Transcription regulation in *Plasmodium falciparum*: Functional characterisation of general transcription factor IIB

Gertrud Talvik

Dissertation presented for the degree of

Master of Science

In the Department of Molecular and Cell Biology

University of Cape Town

February 2016

Supervisor

Dr. Thomas Oelgeschläger

Acknowledgements

I am fortunate to be able to say that the past year has been hard, because hard makes you grow and learn things you otherwise wouldn't have. And boy, I've learnt a lot.

The one who shared the professional journey with me - Robert Milton - I am truly grateful for your support and everlasting faith.
Sara Wighard- thank you for surviving in my blast zone. You are capable of mustering together a solid motivation amidsts utter desperation.
TJ- we've shared this path for five years and sometimes you know me better than I do. Thank you for your (harsh) reality checks and insights.

Ema ja Ats - ma loodan, et see ei jää mu viimaseks ponnistuseks teaduses ja ma saan teile ikka edaspidi kurta kui täbaras olukorras ma jälle olen. Tegelikult oli see üks siiani parimaid kogemusi- nii akadeemilise kui ka isikliku arengu tasandil. Niiet, ema- see oli minu silmis siiani üks su parimaid investeeringuid. Aitäh et te minu vingumisele vastu pidasite, mind toitsite, šokolaadi saatsite ja kaasa elasite. Aitäh Ats, et sa mu teadustööst aru püüdsid saada ja et sinuga sai transkriptsioonist rääkida. Sa olid üks käputäiest inimestest, kes üldse vaevusid.
Ema, Ats, Fränki ja muidugi (eriti) Tontu- ma pühendan selle magistritöö teile.

Isa- sa teadsid juba ammu, et ma suudan rohkemat kui Eesti ärikool. Nüüd ma tean. Aitäh, et sa mind toetasid.
Kollektiivselt mu sponsorid- vantsi ja vantsu, Ester, Sass, Heiki, Rivo- teie toetus oli suureks abiks ja ma olen üdini tänulik.

Craig, you got the best and the worst of it, and I love you for that.

Finally, thank you Dr. Thomas Oelgeschläger for telling me to take life by the balls. I promise I always will. Your guidance and advice was instrumental in this research as well as my future endeavours.

Abstract

Plasmodium falciparum is the causative agent of the most severe form of malaria and continues to pose challenges to international healthcare, with high mortality rates and emergence of drug-resistant strains. *Plasmodium falciparum* has multiple sexual and asexual lifecycle stages within its *Anopheline* mosquito and human hosts, accompanied by distinct morphological changes. The complex lifecycle, along with the ability to adjust rapidly to different environmental niches, is governed by highly regulated and tightly synchronised changes in stage-specific gene expression. In eukaryotes, regulation of RNA Polymerase II transcription initiation is one of the main mechanism of gene expression control. Past research has revealed the presence of crucial elements of RNA Polymerase II (RNAPII) transcription machinery in *P. falciparum*, however, the precise transcription initiation mechanisms in *P. falciparum* remain to date undescribed. Bioinformatics studies have found very little homology between human and *P. falciparum* transcription factors. Furthermore, because of the extreme bias toward A/T in the *Plasmodium* genome, TATA-box and other core promoter elements that direct transcription initiation, cannot be determined with confidence in bioinformatics studies. Functional characterisation of the key protein factors involved in transcription initiation is a first important step towards the identification of core promoter elements and could reveal currently unknown eukaryotic transcription initiation mechanisms or new anti-malarial targets. In eukaryotes, TATA-binding protein (TBP) and transcription factor IIB (TFIIB) are the key protein factors involved in the core promoter recognition and RNAPII pre-initiation complex (PIC) assembly. TBP nucleates PIC assembly by binding the TATA box, thereafter the TBP-TATA complex is further stabilised by TFIIB. In addition, TFIIB has a crucial role in RNAPII recruitment and transcription start site selection and is therefore deemed indispensable in eukaryotic transcription. *P. falciparum* TBP is the only PIC component in *P. falciparum* that has been functionally characterised to date, albeit to a very limited extent. This research study reports the successful expression and purification, as well as initial characterisation of DNA-binding activity of *P. falciparum* TBP and TFIIB. We observe PfTBP binding at multiple locations on putative *P. falciparum* promoters. We further report PfTBP-independent binding activity of PfTFIIB, that has not been previously observed under the conditions and has implications for novel DNA-binding mechanisms of PfTFIIB. Furthermore, we conclusively demonstrate the formation of a PfTBP-PfTFIIB-promoter ternary complex.

Common Abbreviations

A,T,G,C - adenine, thymine, guanine, cytosine
Ad2ML - adenovirus 2 major late
ApiAP2 - Apicomplexan AP2
bp – base pair
BREu/d - upstream/downstream B- recognition element
BSA - bovine serum albumin
C-;N- terminus – carboxyl-; amino- terminus
CAP - catabolite activator protein
cDNA - complementary DNA
DNA - deoxyribonucleic acid
DPE - downstream promoter element
DTT - Dithiothreitol
E. coli - Escherichia coli
EDTA - Ethylenediaminetetraacetic acid
EMSA - electrophoretic mobility shift assay
EtBR - ethidium bromide
GBP-130 - glycophorin binding protein 130
GSH - glutathione
GST - glutathione-S-transferase
GTF - general transcription factor
H. sapiens - *Homo sapiens*
HEPES - 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HRP - horseradish peroxidase
HTH- helix-turn-helix
Inr – initiator
IPTG - Isopropyl β-D-1-thiogalactopyranoside
ITA - immobilised template assay
iRBCs - infected red blood cells
KAHRP - knob-associated histidine rich protein
LB - lysogeny broth
MTE - motif ten element
MW - molecular weight
Ni-beads - Nickel magnetic beads
Ni-NTA - Nickel nitriloacetic acid
OD - optical density
ORF - open reading frame
P. falciparum - *Plasmodium falciparum*
PCR - polymerase chain reaction
PfEMP - *P. falciparum* erythrocyte membrane protein
PfTBP - *P. falciparum* TATA-box binding protein
PfTFIIA - *P. falciparum* transcription factor IIA
PfTFIIB - *P. falciparum* transcription factor IIB
PIC - pre-initiation complex
PVDF - polyvinylidene difluoride

RBC - red blood cells
RNA - ribonucleic acid
RNAPII - RNA polymerase II
S. cerevisiae - *Saccharomyces cerevisiae*
SDS - sodium dodecyl sulphate
SDS-PAGE - SDS polyacrylamide gel electrophoresis
SOC - Super Optimal broth with Catabolite repression
TBE - tris base, boric acid and EDTA
TBP - TATA-box binding protein
TE - Tris-EDTA
TF - transcription factor
TFII- A,B,D - transcription factor II – A,B,D, TAF - TBP-associated factor
TSS - transcription start site
6His- hexa-histidine

Contents

1. Introduction

1.1. The impact of malaria

Malaria is the leading cause of morbidity and mortality worldwide, and continues to pose challenges to international healthcare with the recent emergence of drug-resistant strains. Currently, over 40% of the world's population is at risk of contracting malaria, with over 200 million cases of infection and over 400 000 reported deaths reported in 2015 (World Malaria Report 2015, World Health Organisation 2015). 90% of malaria infections and deaths occur in the African region, with malaria remaining one of the leading causes of child mortality in sub-Saharan Africa. Malaria is caused by parasitic protozoans of the genus *Plasmodium* and transmitted to humans by Anopheles mosquitoes. Four species of *Plasmodium* cause malaria in humans, including *P. vivax*, *P. ovale*, *P. malariae* and *P. falciparum*. The most severe disease progression and the highest mortality rates are caused by *P. falciparum*.

1.2. *Plasmodium falciparum*

1.2.1. Life cycle of *P. falciparum*

P. falciparum is an obligatory parasite belonging to the phylum Apicomplexa. Most Apicomplexa have a complex lifecycle with sexual and asexual stages in multiple hosts. *P. falciparum* parasitism is limited exclusively to its human and female *Anopheles* mosquito hosts. Transmission to humans begins with a bite from an infected female *Anopheles* mosquito, whereby sporozoites from the mosquito salivary glands are released to the bloodstream. Sporozoites rapidly invade the liver cells (hepatocytes), commencing the exo-erythrocytic cycle. In hepatocytes, sporozoites differentiate into trophozoites, which form schizonts by nuclear division into many small nuclei. The multinucleated schizonts undergo multiple fission to form merozoites. Merozoites are passed into the bloodstream, where they access red blood cells (erythrocytes; RBCs), marking the beginning of the asexual intra-erythrocytic stage of the parasite (Cowman and Crabb, 2006). Merozoite invasion of erythrocytes in an essential step in the parasite's life cycle and the erythrocytic cycle is the only stage with manifestation of clinical

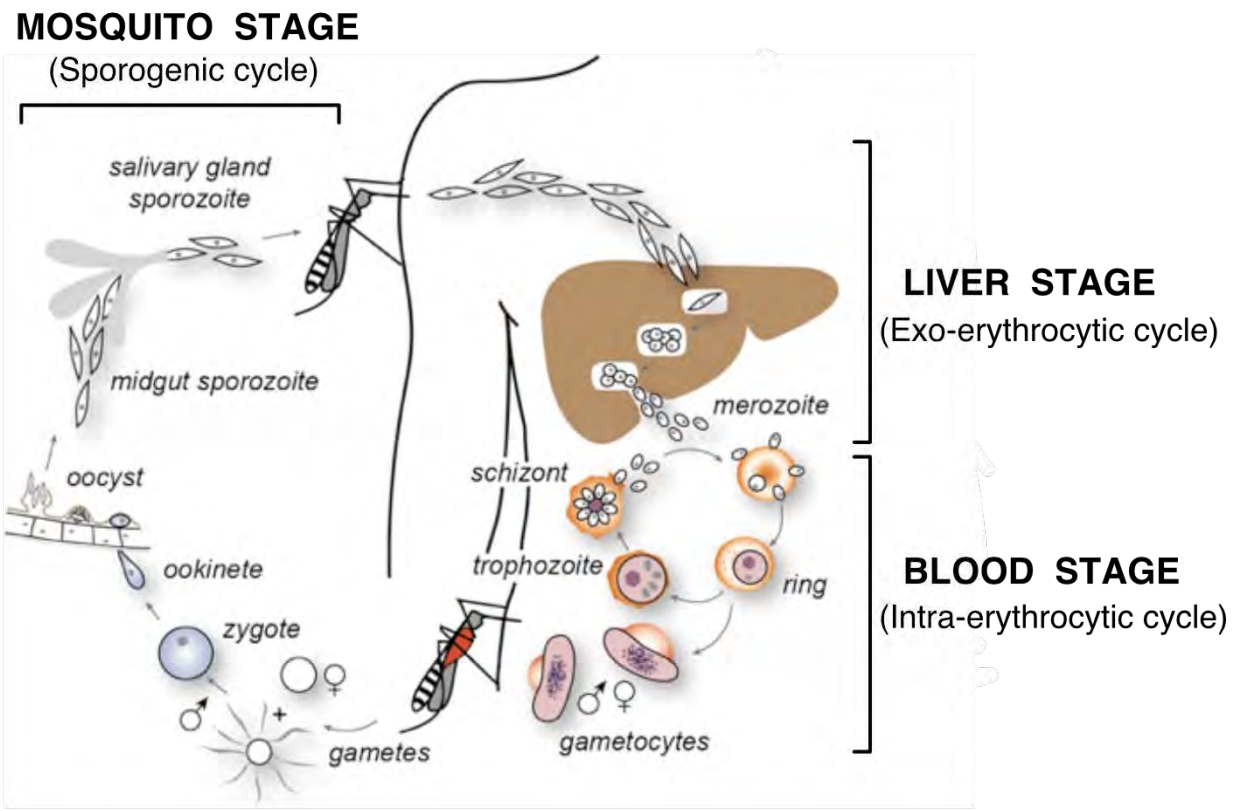

Figure 1. Life cycle of *P. falciparum.*
Schematic of the life cycle stages of *P. falciparum*. Illustration adapted from Cowman et al. (2012)**.**

---

symptoms of disease in the human host. In erythrocytes, the parasite development continues through ring, trophozoite and schizont stages, producing a large number of merozoites that are subsequently released, which proceed to invade other erythrocytes to continue the asexual life cycle (Cowman and Crabb, 2006). Merozoite release from the iRBCs corresponds to intermittent onset of fever (paroxysms) associated with malaria. During the 48-hour intra-erythrocytic cycle, parasite population in the host drastically increases, with up to 50% of RBCs infected upon maximal *P. falciparum* parasitemia.

During the erythrocytic stage, some merozoites develop into male or female gametocytes, that are taken up by a mosquito during the blood meal and thus complete the infectious cycle (Alano, 2007). In the mosquito, the sexual sporogenic cycle of the parasite begins with the formation of zygotes that develop into a mobile ookinete which can traverse through the gut wall to form an oocyst. Sporozoites are released upon the rupture of the oocyst and migrate to the salivary gland of the mosquito for the subsequent infection of humans.

1.2.2. Pathophysiology of *P. falciparum*

P. falciparum infection induces the most severe progression of malaria due to distinct differences in its biology compared to other *Plasmodium* species. *P. falciparum* has developed multiple red blood cell invasion mechanisms and utilises antigenic variation, making it capable of invading a large population of RBCs and evading the host immune system (Miller et al., 2002). A key determinant of the pathophysiology of the parasite is the ability to alter the properties of the infected red blood cells (iRBCs), which as a result adhere to endothelial and other intravascular cells in a process termed cytoadherence. Cytoadherence of the iRBCs is caused by parasite encoded proteins that are presented on the erythrocytic membrane and interact with the surrounding cells. *P. falciparum* systematically alters the epitopes that are exposed to the host immune system, in effect avoiding detection by antibodies and thus ensuring persistent infection (Kirkman and Deitsch, 2012). The primary *P. falciparum* antigen is *P. falciparum* erythrocyte membrane protein one (PfEMP1), which associates with the erythrocyte membrane proteins via *P. falciparum* knob-associated histidine-rich protein (KAHRP). PfEMP1 is expressed in highly variable forms by the *var* gene family (Baruch et al., 1995). *P. falciparum* genome harbours approximately 60 *var* genes, with only one PfEMP1 variant expressed at a time (Scherf et al., 2008). Switching the expression of *var* genes at low frequencies results in antigenically distinct populations of the parasite, which minimises comprehensive immune recognition by the host and ensures persistent infection (Miller et al., 2002). The adhesion of iRBCs to uninfected RBCs and endothelial cells within various organs excludes iRBCs from systemic circulation allowing the parasite to avoid clearance by the spleen. Many PfEMP1 adhesion domains are common among PfEMP1 variants, whereas some variants have specialised to bind host cell receptors within specific organs, such as the placenta or the brain (Fried and Duffy, 1996, Montgomery et al., 2007). Parasite sequestration within vital organs results in serious physiological complications for the human host (Kirkman and Deitsch, 2012).

Highly coordinated and tightly regulated gene expression is a fundamental aspect of the complex life cycle of *P. falciparum* and its sophisticated host immune evasion mechanism. Consequently, understanding the precise mechanisms governing *P. falciparum* gene regulation is an important prerequisite to fully understand *P. falciparum* biology and pathogenesis.

1.2.3. The genome and gene expression of *P. falciparum*

The *P. falciparum* genome is unique in its adenine and thymine nucleotide (A/T) content, accounting for more than 80% of the genome. This extremely A/T-rich genome is unparalleled even within *Plasmodium* species, with the *P. falciparum* genome being the most A/T-rich genome sequenced to date (Gardner et al., 2002). Out of the ~5300 proteins encoded by the genome, 60% remain without a functional assignment due to little or no sequence similarity with other eukaryotes.

Past research has provided evidence for control of gene expression regulation in *P. falciparum* at transcription initiation and post-transcriptional levels, as well as via epigenetic regulation mechanisms such as chromatin remodelling and histone modifications (reviewed in Horrocks et al., 2009). It is likely that a dynamic interplay between multiple levels of gene expression regulation facilitates temporal gene expression in the complex life cycle of the parasite.

Seminal studies into the control of gene regulation on the level of transcription initiation revealed that key eukaryotic features for the transcription regulation of protein-encoding genes are in fact present in *P. falciparum* (Lanzer, 1992a, Lanzer, 1992b). In eukaryotes, a bipartite promoter provides a platform for the assembly and regulation of the RNA Polymerase II (RNAPII) transcription machinery. General transcription factors (GTFs) and RNAPII assemble at the core promoter region, forming a pre-initiation complex (PIC; Figure 2). Regulatory regions of the promoter, termed *cis*-acting elements, are located typically upstream of the core promoter and associate with specific *trans*-acting transcription factors (TFs). These sequence-specific TFs can stimulate or repress the transcriptional activity of RNAPII through regulation of PIC assembly and/or function.

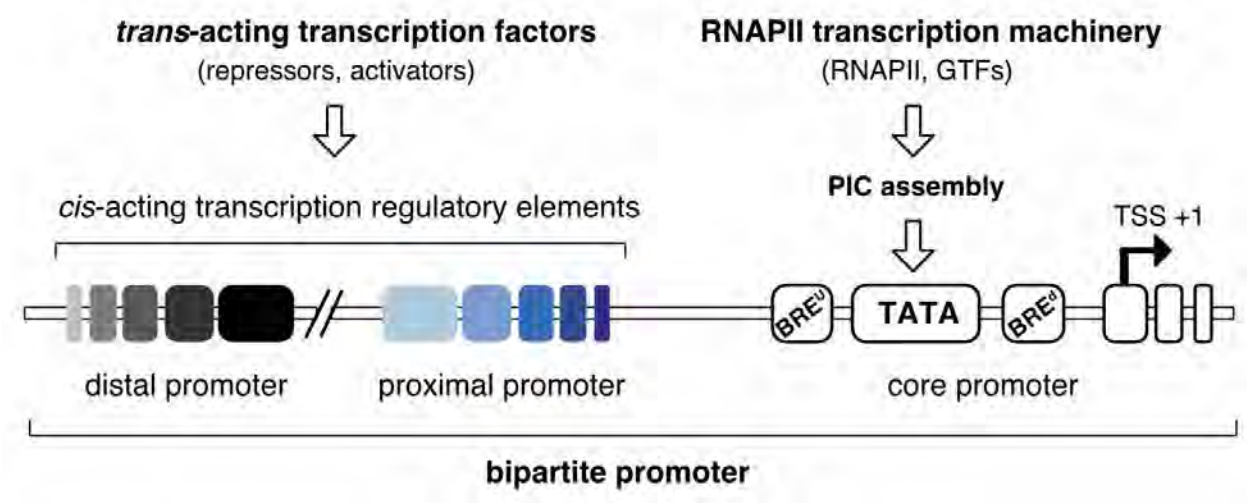

Figure 2. Eukaryotic bipartite promoter
Simplified schematic of the eukaryotic bipartite promoter comprised of regulatory DNA regions (distal and proximal promoter), containing *cis*-acting elements and the core promoter, containing core promoter elements. General transcription factors (GTFs) bind to core promoter elements, such as the TATA box (recognised by TBP) and BREs (BREu and BREd recognised by TFIIB), to form a pre-initiation complex (PIC) with RNAPII. DNA-binding transcription factors (*trans*-acting factors) bind to *cis*-acting elements to regulate PIC assembly and function. TSS, transcription start site (+1).

In *P. falciparum*, association of nuclear *trans*-acting factors with promoter elements upstream from mapped transcription start sites (TSSs) has been shown to correlate with promoter activity and mRNA accumulation (Crabb and Cowman, 1996, Horrocks et al., 2009). Prior to the sequencing of the *P. falciparum* genome, the largest subunit of RNAPII (Li et al., 1989) and a putative orthologue of TATA-binding protein (TBP) - a central component of PIC assembly - were identified and cloned (McAndrew et al., 1993). Further bioinformatics studies utilising hidden Markov profile searching of the *P. falciparum* genome have identified the putative *P. falciparum* orthologues of the remaining, but not all, common eukaryotic PIC components. These include TBP- associated factors (TAFs) that complex with TBP, collectively referred to as transcription factor IID (TFIID), as well as *P. falciparum* orthologues for the transcription factors (TFII-) A, B, B-like, E, F, H and TBP-like (TLP). Interestingly, only three TAFs (1, 2 and 7) have been identified, in comparison to 8-14 TAFs in eukaryotic model systems such as humans, yeast and *Drosophila* (Kokubo, 1994, Burley and Roeder, 1996). With the exception of TAF5, all the TAFs currently unidentified in *P. falciparum* share a conserved histone-fold domain that mediates specific TAF-TAF interactions within the TFIID complex in eukaryotes (Gangloff et al., 2001). In eukaryotes, TAFs 1,2 and 7 interact with TBP independent of other

TAFs (Chalkley and Verrijzer, 1999). Thus, it is likely that the *P. falciparum* TFIID functions in a more minimal form compared to other eukaryotes.

Overall, a substantially smaller proportion of the *P. falciparum* genome is predicted to encode TFs compared to other eukaryotes (1.3% and 4-7%, respectively) (Coulson et al., 2004). The high divergence of the *P. falciparum* genome, exacerbated by the extremely high A/T-content, poses formidable challenges in bioinformatics analyses. Thus, inherent difficulties of *in silico* identification could, at least in part, account for the apparent paucity of *P. falciparum* TFs.

To date, only a few of the *P. falciparum* putative proteins that are predicted to be involved in transcriptional regulation have been characterised *in vitro* or *in vivo*. A family of novel Apicomplexan - specific transcription factors, ApiAP2, have been characterised *in vitro* and shown to bind to specific DNA elements (reviewed in Painter et al., 2011). Five of the ApiAP2 factors have also been characterised *in vivo*, demonstrating involvement in *var* gene silencing and sporozoite development, among other functions (De Silva et al., 2008, Yuda et al., 2010, Painter et al., 2011, Sinha et al., 2014). Two additional sequence-specific TFs, PfMyb1 and PREBP, have been partially characterised and appear to regulate a diverse set of genes modulating the cellular environment and cell cycle progression during the intra-erythrocytic cycle (Gissot, 2005, Komaki-Yasuda et al., 2013).

With regards to the general transcription factors involved in PIC assembly, only the *P. falciparum* putative TBP orthologue has been studied thus far, with preliminary analysis reporting specific interactions between PfTBP and DNA sequences on putative *P. falciparum* promoters (Ruvalcaba-Salazar et al., 2005). Since the stepwise progression towards functional RNAPII PIC assembly is the principal target of transcription regulators in eukaryotes (Roeder, 1998, Hahn and Young, 2011), functional characterisation of other *P. falciparum* GTFs involved in PIC assembly is crucial for understanding transcriptional regulation in the parasite. Importantly, considering the apparent evolutionary divergence between computationally identified *P. falciparum* PIC components and the eukaryotic counterparts (Callebaut et al., 2005, Bischoff and Vaquero, 2010), transcription initiation mechanisms in *P. falciparum* could significantly differ from that of the well-characterised eukaryotic model organisms. Revelation of differences in crucial PIC components between *P. falciparum* and humans could provide new

candidates for anti-malarial drug targets.

Identification of the core promoter elements and *cis*-acting regulatory elements in this extremely A/T- rich genome has been equally challenging. Transcription start sites (TSS) in *P. falciparum* are to date poorly defined (Wakaguri et al., 2009), which further complicates the identification of candidate promoters and regulatory elements *in silico*. Despite numerous *cis*-regulatory elements reported by *in silico* studies, the identity of these sequences remains unclear due to prevailing inconsistencies between various *in silico* approaches (Gunasekera, 2007, Young et al., 2008, Horrocks et al., 2009).

1.2.4. RNA Polymerase II pre-inititation complex

The first step to PIC assembly is TFIID binding to the core promoter. TFIID is a multiprotein complex, consisting of TBP and TAFs. TBP specifically recognises a consensus motif TATAWAAR (in metazoans), whereas TAFs interact with other core promoter elements, collectively conferring TFIID the ability to bind TATA-containing as well as TATA-less promoters (Kim et al., 2005, reviewed in, Thomas and Chiang, 2006). TBP binding nucleates the stepwise assembly of a functional PIC. TBP binding is an induced fit mechanism, as TBP induces a severe bend in the promoter DNA upon binding (Kim et al., 1993, Nikolov et al., 1995). TFIIA is then incorporated, enhancing TBP binding by competing with inhibitory proteins and disallowing TBP homodimerisation (Kokubo et al., 1998, Coleman et al., 1999). TFIIB decreases the dissociation of TBP from the promoter region, effectively enhancing TBP binding to TATA box by stabilisation the TBP-TATA complex (Imbalzano, 1994, Nikolov et al., 1995, Zhao and Herr, 2002). The TFIIB amino terminal domain recruits the RNAPII/TFIIF complex to the promoter and its carboxy-terminal domain exclusively interacts with TFIIB-recognition elements located upstream and downstream of the TATA-box, providing additional support for the TBP-TATA box complex (Nikolov et al., 1995, Lagrange et al., 1998, Deng and Roberts, 2005). TFIIE and TFIIH are then recruited to the promoter and complete PIC assembly. GTFs are also involved in the subsequent interactions with regulatory factors, promoter melting and transcription start site (TSS) recognition (Hahn 2004; Thomas & Chiang 2006).

DNA sequences of the core promoter- including TATA box, the initiator (Inr), downstream promoter element (DPE), motif ten element (MTE) and downstream core element (DCE)- facilitate the PIC assembly, ensure its functional orientation and interact with TAFs (reviewed in Thomas & Chiang, 2006). The combination of core promoter elements differs between genes. TFIIB can make sequence-specific contacts with core promoter elements flanking the TATA-box, termed upstream and downstream IIB recognition elements, BREu and BREd respectively. BREs stabilise the TBP-TFIIB-promoter complex by providing additional DNA contacts, but could also interact with (inhibitory) transcriptional regulators (e.g Mot1, NC2) (Chen and Manley, 2003, Deng and Roberts, 2005).

TATA-binding protein (TBP)

The TATA-box binding subunit of TFIID, TBP, initiates the PIC assembly in eukaryotes. TBP consist of a phylogenetically conserved carboxyl- (C-) terminal domain and a highly divergent amino- (N-) terminal domain. The DNA-binding activity of TBP resides in the C-terminal core domain and is sufficient for PIC formation (Nikolov et al., 1995). The core domain is composed of ten antiparallel β-strands and four α-helices, that form a saddle-shaped molecule. TBP interacts with DNA via the β-strands located in the concave surface and with other transcription

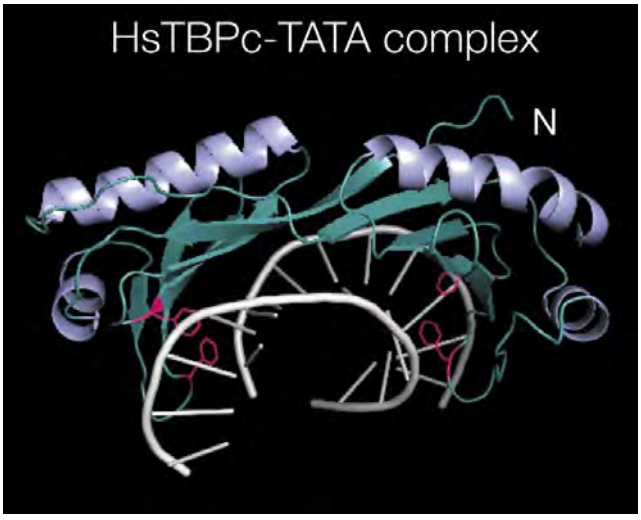

Figure 3. Human TBP core domain binding to the TATA-motif. Tertiary structure of *Homo sapiens* TBP (HsTBP) interacting with the Ad2ML TATA motif (white) via 10 antiparallel β-strands (teal) and four Phe residues (pink). The N- terminus (N) is indicated. Structural data retrieved from Protein Data Bank (PDB) entry 1CDW (Nikolov et al. 1996) and modelled using PyMOL.

factors via the α-helices of the convex surface (Kim et al., 1993, Nikolov and Burley, 1994). TBP distorts and bends the DNA by inserting two phenylalanine (Phe) residues at the 5' and 3'end of the TATA box. Crystallographic studies have concluded that the structure of TBP-DNA complexes is highly similar between species despite of variations in TBP and DNA sequences (Patikoglou, 1999).

The precise function of the N-terminal domain of TBP remains to be assigned. In the absence of an N-terminus, TBP DNA-binding and bending is enhanced (Horikoshi et al., 1990), the effect of transcriptional repressor NC2 is reduced (Goppelt and Meisterernst, 1996) and TBP is more likely to dimerise (Campbell, 2000). Wether these properties have a physiological relevance is yet to be determined.

Comparative bioinformatics studies suggest that the characteristic saddle C-terminal DNA-binding domain of TBP is conserved in *P. falciparum* TBP (PfTBP), despite low primary sequence homology to other eukaryotes (Buendía-Orozco, 2005). *P. falciparum* TBP has been studied previously by Ruvalcaba- Salazar et al. (2005). In this study, the binding activity of recombinantly expressed glutathione S-transferase- (GST-) tagged PfTBP on putative *P. falciparum* promoter regions are investigated using electrophoretic mobility shift assays (EMSAs) and DNAse I footprinting assays. The study reported specific PfTBP binding on TATA-like sequences located 81 and 186 bps upstream of the respective TSS on the tested promoters. The TSS in metazoa is typically located ~25-30bps downstream of the TBP-TATA nucleoprotein complex (Juven-Gershon and Kadonaga, 2010). The results of Ruvalcaba- Salazar et al. (2005) support a TSS selection mechanism similar to *S. cerevisiae*, where transcription initiation occurs ~50-120bps downstream of the TATA element (Zhang and Dietrich, 2005).
Further biochemical studies on PfTBP have not been conducted to date.

Transcription factor IIB (TFIIB)

TFIIB stabilises promoter- bound TBP through protein-protein contacts and interactions with DNA sequences flanking the TATA-box. In addition to its role in PIC assembly, biochemical and structural studies have shown a critical role of TFIIB in TSS selection, promoter escape and initial transcript stabilisation (Tran and Gralla, 2008, Kostrewa et al., 2009, Liu et al., 2010).

TFIIB comprises of four main structural domains. The stable C-terminal core domain (B-core) is composed of two cyclin-like repeats of 5 α-helices (Deng and Roberts, 2007). The aliphatic α-helix situated between at the N-terminal side of the first direct repeat is termed the B-linker and stimulates the catalytic activity of RNAPII (Kostrewa et al., 2009). The N-terminal domain contains a zinc-finger domain (B- ribbon) (Chen et al., 2000), that interacts with RNAPII subunits and TFIIF (Bushnell et al., 2004, Chen and Hahn, 2004) and is conserved in eukaryotic and archaeal TFIIBs (Qureshi and Jackson, 1998, Gietl et al., 2014). Adjacent to B-ribbon is a a helix-loop structure termed B-reader that interacts with RNAPII and is involved in TSS localisation (Kostrewa et al., 2009, Grunberg and Hahn, 2013).

TFIIB is known to enhance TBP binding to the TATA motif by stabilising the bent TBP-TATA complex through direct interactions with the convex side of TBP (Zhao and Herr, 2002). Besides TFIIB-TBP interaction, first analysis of the crystal structure of the ternary complex of TBP and TFIIB core domains (TBPc and TFIIBc, respectively) with a TATA-consensus sequence, indicated that TFIIBc contacts the DNA major groove directly upstream, and the DNA minor groove directly downstream of the TATA motif (Nikolov et al., 1995). A consensus sequence 5'-G/C-G/C-G/A-C-G-C-C-3' was identified that interacts with TFIIB in a TBP-independent

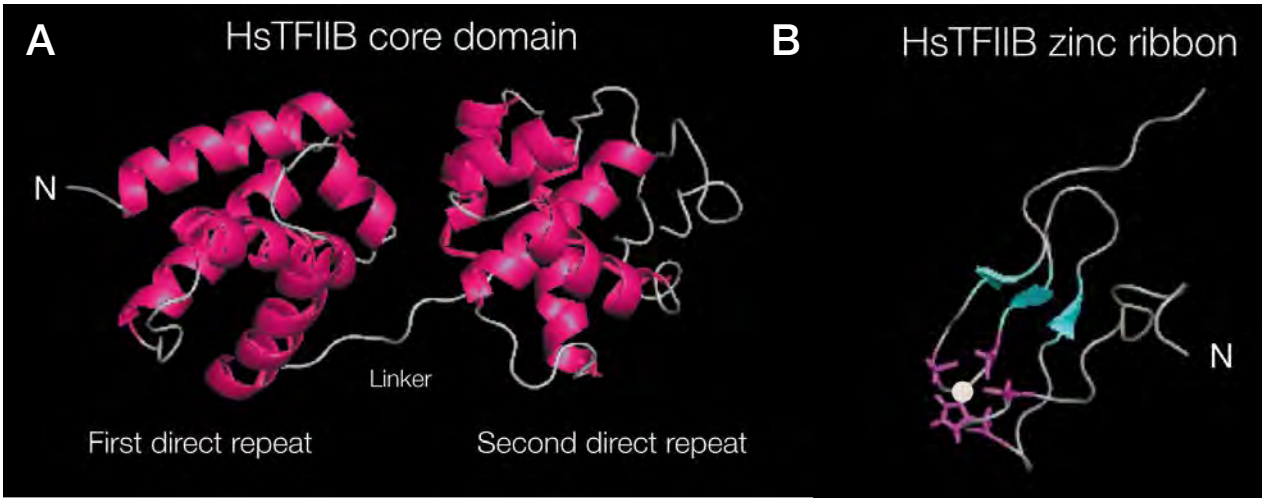

Figure 4. Tertiary structure of human TFIIB core domain and N-terminal zinc ribbon.
A. *Homo sapiens* TFIIB (HsTFIIB) C-terminal core domain consisting of two cyclin-like repeats of α-helices (pink), separated by a linker (indicated). The N-terminus (N) is indicated.
B. HsTFIIB N-terminal zinc ribbon motif. β-strands are indicated by arrows (teal). The zinc (gray sphere) and cysteine/histidine residues that coordinate the ion are shown (pink). The N-terminus (N) indicated.
Structural data retrieved from Protein Data Bank (PDB) entry 1VOL (A; Nikolov et al. 1995) and 1DL6 (B; Chen et al., 2000) and modelled using PyMOL.

manner (Lagrange et al., 1998). TFIIB contacts the BREu via a helix-turn-helix (HTH) motif, a trihelical bundle in the second direct repeat of the C-terminus (Lagrange et al., 1998). Further binding-site selection studies identified a downstream TFIIB recognition element (BREd), with a consensus sequence 5'-G/A-T-T/G/A-T/G-G/ T-T/G-T/G-3' (Deng and Roberts, 2005). This interaction was only observed in the presence of TBP, indicating a the necessity for the severe bend in the DNA for BREd to become accessible to TFIIB (Lagrange et al., 1998, Deng and Roberts, 2005). BREu and BREd are found with similar prevalence in TATA-containing, as well as TATA-less promoters and the conservation of BREs between species is modest (Deng and Roberts, 2005, Gershenzon and Ioshikhes, 2005).

TFIIB N-terminal interactions with RPB1, RPB2 and RPB9 subunits of RNAPII facilitate the recruitment RNAPII/TFIIF to the PIC (Bushnell et al., 2004, Chen and Hahn, 2004). TFIIB is thought to allosterically reconfigure the RNAPII active site and thus stimulate its catalytic power (Tran and Gralla, 2008, Kostrewa et al., 2009). A number of studies have shown that mutations in the B-reader helix shift the TSS and mutations in the B-reader loop impair DNA opening around the TSS (Zhang et al., 2002, Tran and Gralla, 2008). It is suggested, that the TSS is recognised at least partially by TFIIB through the identification of Inr motifs, whereas RNAPII itself might have affinity for initiation at particular sequences (Fishburn and Hahn, 2012).

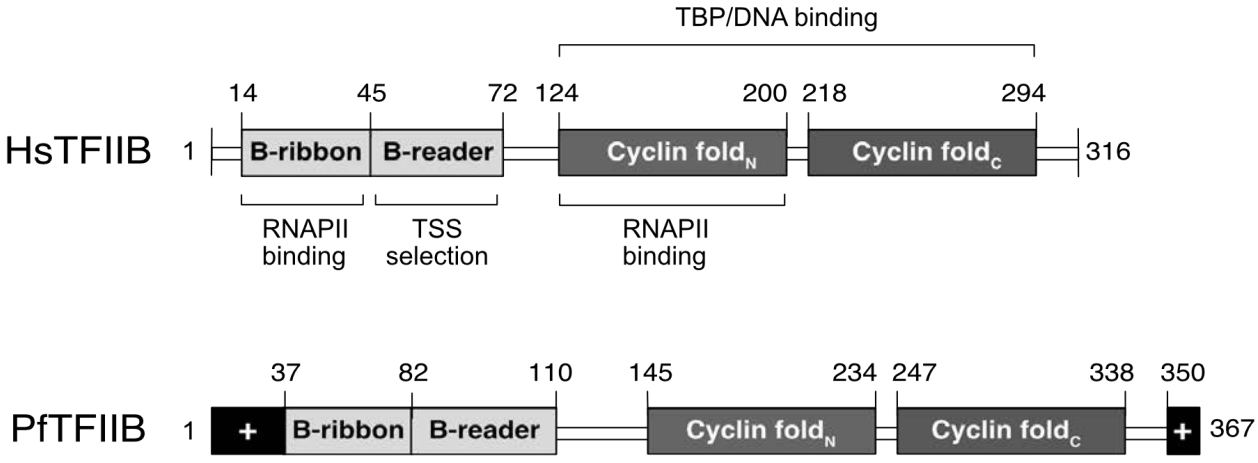

Figure 5. Comparison of human and *P. falciparum* TFIIB main structural features.
The amino acid positions of secondary structures are indicated. PfTFIIB contains basic regions (+) spanning amino acids 1-37 (pI 11.8) and 350-367 (pI 10.7). Main functional interactions of HsTFIIB domains are indicated.

The TFIIB orthologue in *P. falciparum* has been identified (Bischoff and Vaquero, 2010), but remains to be functionally characterised. Previous analyses within our research group suggested a conservation of the main structural features of PfTFIIB compared to eukaryotic TFIIBs despite low amino-acid sequence identity (Steven Bing, MSc thesis 2015). However, some *P. falciparum* specific regions were identified, most notably a highly basic extreme N-terminal region, absent in other eukaryotes and *Plasmodium* species. The TBP-interaction surfaces of PfTFIIB were found to be reasonably conserved. The BREd-binding region in the C-terminal core domain was predicted to be reasonably conserved, whereas similarly to plants and fungi (Lagrange et al., 1998), the HTH motif that contacts BREu was found to be largely absent (Steven Bing, MSc thesis 2015).

1.3. Objectives of this study

Because of the extreme bias towards A/T base pairs in the *P. falciparum* genome, TATA-box and other core promoter elements that direct transcription initiation remain to be determined with confidence in bioinformatics studies. However, the identification of core promoter regions is a crucial first step towards studying transcription regulation. Our laboratory has begun to identify and to functionally characterise the key components of the *P. falciparum* transcription machinery responsible for core promoter recognition and PIC assembly. The objective is to examine the DNA-binding characteristics of GTFs in order to identify *P. falciparum* core promoter regions and the DNA elements directing transcription initiation complex assembly.

In this MSc research work, I aim to investigate the function of TFIIB in the first stages of *P. falciparum* PIC assembly. This investigation includes the cloning and recombinant expression of PfTBP and PfTFIIB, and establishment of a high standard purification scheme to ensure high purity of the proteins of interest for DNA-binding assays. I aim to assess the potential DNA-binding properties of PfTBP and PfTFIIB on putative *P. falciparum* as well as eukaryotic model promoters. Overall, this study aims to provide first insights into the DNA sequence determinants that direct PfTBP/PfTFIIB promoter binding.

2. Materials and Methods

2.1. Molecular cloning techniques

Vector information

The plasmid vectors constructed in this study are derived from commercially available pET11d vectors (*Novagen*). Detailed plasmid maps of the final vector constructs can be found in Appendix A. pET11d-GST-6His-PfTLP and pET11d-6His-PfTBP vectors for cloning purposes were generously provided by Steven Bing and Dr. Thomas Oelgeschläger, respectively.

Primer information

The primers used for cloning are outlined in Appendix B: Table 1. Relevant properties are indicated.

Polymerase chain reaction (PCR)

In all instances, PCR reactions were performed using Kapa Taq polymerase in conjunction with buffer A (1.5mM Mg^{2+}) (*Kapa Biosystems*) in a standard PCR protocol (Sambrook and Russell, 2001), unless stated otherwise.

The construction of pET11d-6His-PfTFIIB-GST included the generation of an intermediate vector, pET11d-6His-GST from a pre-existing pET11d-6His-PfTBP vector. Briefly, a GST ORF flanked by appropriate RE sites was amplified using two individual PCR fragments (amplified from pET11d-GST-6His-PfTLP vector using primers 1 & 2 and 3 & 4, see Table 2) in an overlap extension PCR with primers 5 & 6, as described previously (Heckman and Pease, 2007).

The construction of pET11d-6His-PfTFIIB, 6His:PfTFIIBΔ37 & 6His:PfTFIIBΔ140 vectors included the amplification of PfTFIIB ORF from previously constructed pET11d-6His-PfTFIIB-GST with respective primers (Appendix A: Table 1).

Colony PCR for the identification of positive clones were performed as described by Sambrook and Russell (2001).

PfTFIIB open reading frame (ORF) was amplified from MRA-296 *Plasmodium falciparum* 3D7

16hr cDNA bacteriophage (λZapII) library containing the ring stage of the parasite in 50µl reaction volume containing 50µM respective primers (Table 1., primers 7 & 8). In order to efficiently amplify A/T- rich *P. falciparum* sequences, the PCR elongation step was carried out at 60˚C (Su et al., 1996).

All PCR products were purified using either the MinEluteTM Reaction Cleanup Kit (*Qiagen*) or QIAquick PCR Purification Kit (*Qiagen*), following the manufacturer's instructions. The purity and integrity of purified DNA was verified by agarose electrophoresis and DNA concentrations were determined with NanoDrop 2000 spectrophotometer (*Thermo Scientific*).

Table 1. General PCR profile for the amplification of P. falciparum sequences

Duration	Temperature (C°)	Cycles
5min	95	1
30sec	95	30
30sec	56	
1min	60	
10min	60	1
10min	72	1

Agarose gel electrophoresis

DNA was analysed on 0.8-2% (w/v) agarose gels containing 0.02µg/ml ethidium bromide (EtBR) in 0.5×TBE (40mM Tris-Cl pH 8.3, 45mM boric acid, 1mM ethylenediaminetetraacetic acid [EDTA]) buffer. Electrophoresis was carried out at 80V until sufficient separation. In all instances GeneRuler DNA ladder (*Thermo Scientific*) was used for size comparison and estimation of DNA yield.

Restriction enzyme digestion and ligation

6His:PfTFIIB:GST construct

The overlap extension PCR product and pET11d-6His-PfTBP was digested with BamHI (*Thermo Scientific*) and *Eco*RI (*Fermentas*) and ligated. The resulting intermediate vector, pET11d-6His-PfTBP-GST, was digested with *BamH*I (*Thermo Scientific*) and *Nde*I (Thermo Scientific) and PfTFIIB ORF was digested with *Bgl*II (Fermentas) and *Nde*I (*Thermo Scientific*)

and ligated into resulting vector frame.

6His:PfTFIIB, 6His:PfTFIIBΔ37 & 6His:PfTFIIBΔ140 constructs

PfTFIIBΔ37 and PfTFIIBΔ140 PCR products and PfTFIIB ORF was digested with *Bgl*II (Fermentas) and *Nde*I (*Thermo Scientific*). *BamH*I/*Nde*I digested and de-phosphorylated pET11d-6His vector frame was provided by Dr. Thomas Oelgeschläger.

QIAquick MinElute Reaction Cleanup Kit 50 (*Qiagen*) or QIAquick Gel Extraction Kit 50 (*Qiagen*) were used for the cleanup of all restriction enzyme (RE) digestions. Manufacturer's instructions were followed in all procedures.

In all instances, vector frames were dephosphorylated using alkaline phosphatase (*Thermo scientific FastAP*) as specified by the manufacturer. All ligations reactions were carried out at room temperature (RT) for 15-60min using 1:1 and 3:1 insert to vector ratios with T4 DNA ligase (*Thermo Scientific*).

Transformation into E. coli TOP10 cells and plasmid isolation

In all instances, 10µl of ligation mix was transformed into 100µl OneShot® TOP 10 Chemically Competent *E. coli* cells (*Invitrogen*) following manufacturer's instructions and plated on lysogeny broth agar plates (LB; 1% [w/v] tryptone powder, 0.5% [w/v] yeast extract, 0.5% [w/v] NaCl, 1.5% [w/v] agar, pH 7.5) containing 100µg/ml ampicillin. Colonies containing positive clones were identified by colony PCR as described above and used to inoculate 10mL lysogenic broth (LB; 1% [w/v] tryptone, 0.5% [w/v] yeast extract, 170mM NaCl) containing 100µg/ml ampicillin and grown overnight (O/N) at 37˚ C. Plasmid extraction was performed using GeneJET Plasmid Miniprep Kit (*Thermo scientific*) following manufacturer's instructions.

Sequencing

The sequencing was carried out by Stellenbosch Central Analytical Facility (Stellenbosch University, South Africa).

Transformation into E. coli BL21 cells for protein expression

BL21-CodonPlus (DE3)-RIL cells (BL21; *Agilent Technologies*) were used for the recombinant expression of all *P. falciparum* proteins. The BL21-CodonPlus (DE3)-RIL cells contain a chloramphenicol resistant plasmid with extra copies of genes encoding tRNA's for arginine (AGA; AGG), isoleucine (AUA) and leucine (CUA) codons. Respective tRNA's are rare in *E. coli* translation events and supplementation ensures optimal expression of proteins with high frequency of these codons.

All expression vectors used in this study were transformed as described above. Transformants were spread plated onto Super Optimal Broth with Catabolite Repression (SOC; 2% [w/v] tryptone powder, 0.5% [w/v] yeast extract, 10mM NaCl, 2.5mM KCl, 10mM $MgCl_2$, 20mM glucose, 1.5% [w/v] agar) plates containing 50µg/ml chloramphenicol and 100µg/ml ampicillin, and incubated overnight at 37°C.

2.2. Expression of recombinant *P. falciparum* proteins in *E. coli*

All expression cultures listed below contained 50µg/ml chloramphenicol, 100µg/ml ampicillin. In all instances, cell growth was monitored by measuring the optical density of the cultures (OD) with a commercial spectrophotometer.

Starter culture

A single colony of transformed *E. coli* BL21 cells was used to inoculate 10-25mL of LB (1% [w/v] glucose) and incubated O/N at 37˚C. The culture was centrifuged at ~12000 x g for 10min at 4˚C and bacterial pellets re-suspended in fresh LB (1% [w/v] glucose) in order to remove secreted β-lactamase. Expression cultures were inoculated with starter culture in a 1:20 or 1:100 ratio, as described.

Expression of 6His:PfTBPcore

Starter culture containing *E. coli* BL21 cells transformed with pET11d-6His-PfTBP was used to inoculate 500mL LB at a ratio of 1:100 (v/v). Expression cultures were grown at 37˚C until mid-log phase, OD_{550}=0.45 (~3hours). The cultures were cold shocked on ice for 90 seconds and

incubated for a further 16 hours (O/N) at 30˚C. The cold-shock method has been reported to lead to increased expression of cold-shock chaperone proteins (Qing et al., 2004), thus potentially reducing aggregation and improving solubility of recombinant PfTBPc. The cultures were pelleted by centrifugation at ~5500 x g at 4˚C for 10min and flash-frozen in liquid N_2 for storage at -80˚C.

Expression of 6His:PfTFIIB:GST

Starter culture containing *E. coli* BL21 cells transformed with pET11d-6His-PfTFIIB-GST was used to inoculate 500mL LB (1% [w/v] glucose) at a ratio of 1:20. Expression cultures were grown at 37˚C for 5 hours and pelleted by centrifugation at ~12000 x g for 10min at 4˚C. For induced cultures, at OD_{600}=~1.0, 0.1mM isopropyl 1-thio-ß-D-galactopyranoside (IPTG) was added to the culture and grown for another 2.5h. Pellets were obtained as described above and flash-frozen in liquid N_2 for storage at -80˚C.

Expression of 6His:PfTFIIB, 6His:PfTFIIBΔ37 & 6His:PfTFIIBΔ140

Starter culture containing *E. coli* BL21 cells transformed with the pET11d-6His-PfTFIIB, pET11d-6His-PfTFIIBΔ37 or pET11d-6His-PfTFIIBΔ140 was used to inoculate 500mL LB at a ratio of 1:20. Expression cultures were grown at 37˚C for 5 hours and pelleted by centrifugation at ~12000 x g for 10min at 4˚C. Pellets were flash-frozen with in liquid N_2 for storage at -80˚C.

2.3. Purification and analysis recombinant *P. falciparum* proteins

All purification procedures were carried out at 4˚C using gravity flow where applicable. Elution fractions were flash-frozen in liquid N_2 and stored for analysis at -80˚C. BC- buffers (20mM Tris, pH 7.2 at RT, 0.02mM EDTA, 10mM β-mercaptoethanol, 0.1% NP 40) were used for protein storage and purification, with the respective KCl concentration indicated (e. g. BC-100 contains 100mM KCl).

Isolation of soluble proteins from E. coli cellular lysates

The following procedures were carried out on ice. Cell pellets were resuspended in sonication buffer (50mM Tris-HCl, 300mM KCl, 0.2% Nonidet P-40 [NP-40], 10mM ß-mercaptoethanol

[ß-EtSH]) with the addition of protease inhibitor cocktail (50µl/g per cell pellet; P8340 *Sigma-Aldrich*), lysozyme (0.5mg/ml; Sigma-Aldrich) and Benzonase (1µl/5g cell pellet; E1014 *Sigma-Aldrich*). The cell suspension was left on ice for 30min followed by sonicated in 6x10s bursts with 50s cooling on ice between bursts. The resulting cell lysate was centrifuged at 30000 x g for 30min at 4˚C to obtain cleared lysate.

Purification of 6His:PfTBPc by Nickel-affinity chromatography and ion exchange chromatography

6His:PfTBPc was purified from cleared lysate on Nickel-Nitrilotriacetic acid (Ni-NTA) agarose (Nickel Affinity Gel, *Sigma-Aldrich*) in Poly-Prep® Chromatography columns (*Bio-Rad*). Cleared lysate was passed through the column, washed with 20x packed column volume (PCV) of wash buffer (20mM Tris-HCl pH 7.2, 100mM KCl, 10mM ß-EtSH, 0.1% NP-40, 20mM imidazole), and eluted with ~5 PCV of elution buffer (20mM Tris-HCl pH 7.2, 20% (v/v) glycerol, 0.2mM EDTA, 100mM KCl, 10mM ß-EtSH, 0.1% NP-40, 250mM imidazole). 200-600µl elution fractions were collected in sterile eppendorf tubes.

Ni-NTA agarose elutions were analysed by SDS PAGE and 6His:PfTBPc containing fractions pooled and adjusted to 500mM KCl using PD-10 Desalting columns (*GE Healthcare*), according to manufacturer's instructions. PfTBPc was further purified on PureProteome™ nickel magnetic beads (Ni-beads; *Millipore*) as specified by manufacturer, with following adaptions: cleared lysate was incubated with the beads at 4˚C with shaking for 1h and beads separated from the unbound protein fraction using a magnetic separator; Ni-NTA agarose elution buffer was as described above.

Further purification using ion exchange chromatography was carried out by Dr. Thomas Oelgeschläger. Briefly, Ni-NTA elution fractions were pooled, adjusted to 100mM KCl using PD-10 Desalting columns (*GE Healthcare*), and subjected to Sulphopropyl- Sepharose® Fast Flow (SP-Sepharose; *Sigma-Aldrich*) and eluted in a stepwise manner with BC-120 up to BC-280 (20mM KCl steps), followed by BC-300 and BC-500. Protein containing fractions were identified by SDS PAGE, pooled and adjusted to 100mM KCl using PD-MidiTrap G-25

desalting columns (*GE Healthcare*), and passed over Q-Sepharose® Fast Flow (*Sigma Aldrich*). Protein containing fraction was recovered in the flowthrough.

Purification of 6His:PfTFIIB:GST by glutathione and nickel affinity chromatography

Glutathione-agarose (GSH) resin (Sigma-Aldrich) was prepared following manufacturer's instructions. Cleared lysate was subjected to the resin, followed by wash with 20x PCV of wash buffer (20mM Tris-HCl pH 7.2, 0.2mM EDTA, 100mM KCl, 10mM ß-EtSH, 0.1% NP-40) and bound proteins were eluted with glutathione elution buffer (20mM Tris-HCl pH 7.2, 20% (v/v) glycerol 0.2mM EDTA, 100mM KCl, 10mM ß-EtSH, 0.1% NP-40, 20mM glutathione). The GSH eluates were loaded onto Ni-NTA agarose HIS- Select® HF Nickel Affinity Gel (*Sigma-Aldrich*), followed by wash with 20x PCV of wash buffer. The proteins were eluted in NI-NTA elution buffer, as described above.

Purification of 6His:PfTFIIB by nickel affinity chromatography and ion exchange chromatography

Ni-NTA agarose HIS- Select® HF Nickel Affinity Gel (*Sigma-Aldrich*) was prepared as specified by manufacturer and purification procedure followed as described above. Ni-NTA agarose purification elution fractions were pooled and buffer exchanged into fresh BC-100 (20 mM Tris, pH 7.2 at RT, 100mM KCl, 0.02mM EDTA, 10mM β-EtSH, 0.1% NP 40) using PD-10 Desalting column (*GE Healthcare*), according to manufacturer's instructions.
SP-Sepharose (*Sigma-Aldrich*) resin was prepared as specified by the manufacturer and equilibrated in BC-100 buffer. Pooled Ni-NTA agarose elution fractions were passed over SP-Sepharose and flowthough collected. The resin was washed with 5x PCV of BC-100, and bound proteins were stepwise eluted with 4x PCV of BC-buffer, containing 200mM KCl, followed by 350mM and 500mM KCl. ~500µl fractions were collected in sterile eppendorf tubes.

SDS PAGE analysis

In all instances the Mini-PROTEAN® Electrophoresis System (*Bio-rad*) was utilised. 12-15%

resolving and 5% stacking gels were prepared according to standard protocol (Sambrook and Russell, 2001). Samples were mixed in a 1:1 ratio with 2×SDS loading dye (100 mM Tris-Cl (pH 6.8), 4% [w/v] SDS, 0.2% [w/v] bromophenol blue, 20% [v/v] glycerol, 200mM dithiothreitol [DTT]) and heated at 80°C for 2 minutes. Samples were electrophoresed at a constant amperage of 40mA until completion. The gels were washed in deionised water for 3 x 10 min, stained with Bio-Safe^{TM} G-250 Coomassie premixed stain (*Bio-Rad)* for 30 minutes and subsequently destained with deionised water. In all instances, Precision Plus All Blue (*Bio-Rad*) and Color Prestained Protein Standard (*NEB*) molecular weight (MW) markers were used, as indicated in figure legends.

Immunoblot analysis

Immunoblots were performed utilising the Mini Trans-Blot® Electrophoretic Transfer Cell (Bio-Rad) apparatus, following manufacturer's instructions.

Proteins of interest were separated on SDS PAGE and transferred to a Millipore Immobilon-P transfer membrane (PVDF) using a standard protocol (Sambrook and Russell, 2001). The membrane was washed in 1×TBST buffer (137mM NaCl, 2.7mM KCl, 19mM Tris, 1% Tween-20 [*Sigma-Aldrich*]) 3 times for 5 minutes, followed by deionised H_2O. The membrane was stained for 1 hour with Colloidal Gold Total Protein Stain (*Bio-Rad*), followed by wash with deionised H_2O. The membrane was washed 3 times with 1×TBST buffer and blocked with 5% [w/v] fat-free milk in 1×TBST for 30min at RT and incubated with the appropriate antisera/ purified antibody (1:500 and 1:1000 dilution, respectively, in 5% [w/v] fat-free milk in 1×TBST) for 2h at RT or O/N at 4˚C. The membrane was washed with 3 times with 1xTBST and incubated with horseradish peroxidase-conjugated anti-rabbit IgG (1:3000 in 1xTBST; Sigma-Aldrich) for 30min at RT, followed by 3 x wash with 1xTBST. Immunoreactant proteins were visualised with Pierce® ECL Western blotting substrate (Thermo Scientific), according to manufacturer's protocol on X- ray film (CL-XPosure, *Thermo Scientific*).

2.4. Affinity purification of antibodies for PfTBP, PfTFIIB and PfTFIIAα/β

Polyclonal rabbit peptide antibodies against PfTBP (residues 125- 138: C-SEYDNNE KEKSDDL), PfTFIIB (residues 354-367: C-LKQKYLSEDKRKKN) and PfTFIIAα/β (residues:C-GLVSNKKENKNSKI) were produced by BioGenes (Berlin, Germany). The presence of the peptide antibodies in the rabbit antisera was confirmed by ELISA. Further purification of the polyclonal antibodies from rabbit antisera was carried out using antigen-affinity purification. Peptides were coupled to SulfoLink™ Coupling Resin (*Thermo Scientific*) and antigen-affinity purification was carried out following the protocols by the Hyman Lab at the Max Planck institute of Molecular Cell Biology & Genetics (detailed protocol can be accessed at: http://hymanlab.mpi-cbg.de/hyman_lab/), with the following adaptions: the PfTBP, PfTFIIB or PfTFIIAα/β lypholized peptide was dissolved in 50mM Tris, 5mM EDTA, pH 8.5 to 1mg/mL and incubated in the resin for 30min at RT; serum was diluted with an equal volume of 20mM Tris pH 7.5, centrifuged at 20000x g at 4˚C for 30min and decanted into a sterile container; bound antibodies were eluted with 100mM glycine pH2.5 followed by 100mM triethylamine pH11.5. 10µl diluted BioRad Protein Assay Reagent (*BioRad*) was mixed with 50µl elution fractions and antibody-containing fractions were identified and pooled. The pooled fractions were dialyzed against 3 x 2L phosphate-buffered saline (PBS; 137 mM NaCl, 2.7 mM KCl, 10mM Na_2HPO_4, 1.8mM KH_2PO_4) and concentrated using Nanosep® Centrifugal Device with Omega Membrane 10K (*Pall*), according to the manufacturer's instructions. Purified antibodies were analysed on SDS PAGE and protein concentration determined by BioRad Protein Assay (*BioRad*), according to manufacturer's instructions.

2.5. DNA-binding assays

The DNA-binding activities of PfTBP and PfTFIIB are investigated using two independent approaches - electrophoretic mobility shift assays (EMSAs) and immobilised template assays (ITAs).

Electrophoretic mobility shift assays

Electrophoretic mobility shift assay (EMSA) is a well established method of identifying protein-DNA complexes (Garner and Revzin, 1981). The classical method is based on autoradiographic

detection of the mobility of ^{32}P-labelled nucleic acid in a native gel matrix. The nucleic acids are incubated with the protein of interest and upon DNA-binding activity of the proteins the relative mobility of the DNA decreases, resulting in a slower migrating protein-DNA complex, i. e. a 'gel shift' (Hellman and Fried, 2007). Chemiluminescence detection is a common alternative to radiographic detection (Rodgers et al., 2000) and was utilised in this study. In this method, the target DNA 5' and 3' ends are labelled with biotin which interact with streptavidin- conjugated horseradish peroxidase resulting in chemiluminescence upon addition of substrate. The biotinylated, electrophoresed DNA probes are transferred to a positively charged nylon membrane, incubated with the streptavidin conjugate and luminescence is detected by X-ray film. In this study, polyacrylamide gels as well as Mg^{2+} agarose gels were used. Magnesium-containing agarose gels have been used as an alternative to non-denaturing polyacrylamide (Hellman and Fried, 2007), as larger pore size of agarose allows for the resolution of large multi-protein complexes (Zerby and Lieberman, 1997).

Immobilised template assays

Immobilised template assays are commonly used to study protein-DNA interactions and cooperative binding of pre-initiation complexes (Johnson et al., 2004). It provides an alternative for the investigation of large complexes that are often difficult to resolve in EMSAs and provides further evidence for specific protein-DNA interactions. In this assay, DNA probes are 5' biotinylated and immobilised on streptavidin-coated magnetic Dynabeads™. DNA-binding proteins are incubated with the immobilised beads and unbound fraction is removed by separating the immobilised template with a magnetic particle concentrator. The bead pellet is washed to ensure the elimination of non-specific binding and proteins bound to the template DNA are eluted and detected by immunoblotting.

Promoter sequences used in this study

The promoter regions of two *Plasmodium*-specific genes, *kahrp* and *gbp-130*, and the eukaryotic model promoter, adenovirus 2 major late (Ad2ML), were chosen for DNA binding assays. The adenovirus 2 promoter contains a classical TATA-box motif, and has been widely utilised in eukaryotic transcription studies (Thomas and Chiang, 2006). This promoter allows for

comparison with respective human transcription factors and for the identification of potential sequence preference of *Plasmodium* factors.

The *Plasmodium* genes chosen for this study, *kahrp* and *gbp-130*, are involved in mediating the pathogenicity of the parasite in humans. The *gbp-130* gene encodes glycophorin binding protein- an essential mediator which binds erythrocyte binding protein and initiates the invasion of RBCs by the parasite (Perkins, 1984). The *kahrp* gene encodes knob-associated histidine-rich protein, which has been shown to interact with erythrocyte membrane skeleton proteins. Studies link these interactions to cytoadherence of RBCs, which is the main determinant of the pathophysiology of malaria (Pei et al., 2005, Weng et al., 2014).

Few studies have investigated these genes to date. Proteins from *Plasmodium* nuclear extracts have been shown to associate with *gbp-130* and *kahrp* promoter regions in a stage-specific manner, and the TSS have been mapped (Lanzer, 1992a, Lanzer, 1992b). PfTBP binding of the promoter regions of these genes has been demonstrated as by Ruvalcaba-Salazar et al (2005). The aforementioned study mapped PfTBP binding to TATA-like elements in regions located at -81 and -186bp relative to TSS. The probes used in this MSc research were designed to accommodate this observation, spanning approximately +60 to -170 from the reported TSS, resulting in 282bp sequences (see Table 2). The Ad2ML promoter used was 297bp in length. All EMSA DNA probes were 5' and 3' double biotinylated.

The upstream region of highly variable (*var*) genes, was identified as a region of interest in this study and included as DNA target for PfTFIIB in polyacrylamide EMSAs. The *var* genes encode PfEMP1 that confers antigenic variation and is the parasite's main mechanism of evading host immune attack (see Chapter 1) (Kirkman and Deitsch, 2012). The upstream regions of *var* genes are highly conserved, with deletions leading to promoter inactivation in transgenic *P. falciparum* lines (Brancucci et al., 2012). Therefore, a 60bp probe spanning from -80 to -21 relative to TSS, including a conserved 31bp region of var 5B1 identified previously by Voss et al. (2000), was included in this study.

The Ad2ML, *kahrp* and *gbp-130* promoter regions were also used in ITAs (Table 2). In order to avoid steric hindrance, the ITA probes were amplified with five GAL4 binding sites, allowing

for sufficient space between potential DNA binding sites and magnetic beads. The *P. falciparum* ITA probes were therefore longer (422bp) and were biotinylated only at the 5' end.

Preparation of DNA probes used in EMSAs and ITAs

Vectors containing the *Plasmodium* promoter sequences were produced by Tomas Hessler, an Honours research student in the laboratory. In brief, the regions spanning +51 to -181bp relative to the *gbp-130* TSS and +63 to -172 relative to the *kahrp* TSS and were amplified from *Plasmodium* 3D7 genomic DNA library and cloned into a pGEM-7Zf (*Promega*) derived vector. The Ad2ML probe was amplified from the pTOG5ML(-51TATA/+62) vector (Oelgeschlager et al., 1998). The *kahrp*, *gbp-130* and Ad2ML promoter regions were PCR amplified with biotinylated primers and purified using QIAquick Gel Extraction Kit 50 (*Qiagen*).

The *var* probe oligonucleotides were synthesised by the Oligonucleotide Synthesis lab (UCT@MCB, Cape Town). 10µM (+) and (-) strands were hybridised by heating in Tris-EDTA (TE, Tris-HCl pH 7.2, 1mM EDTA) containing 50mM NaCl at 95˚C for 5min and allowed to cool to RT over ~30min. The annealed oligonucleotides were resolved by PAGE and purified using QIAEXII Gel Extraction Kit (*Qiagen*), according to the manufacturer's instructions.

Unbiotinylated oligonucleotide competition probes

Unlabelled full length *kahrp*, *gbp-130* and Ad2ML promoters were amplified as described above using unbiotinylated primers. TATA-containing and TATA-less sequences were amplified from the pTOG5TdT(-41TATA/+33) and pTOG5TdT(-41/+33) vectors (Malecova et al., 2007). Primers for the amplification of *kahrp* promoter regions spanning nucleotides 1-138 (termed *kahrp* A) and 125-282 (termed *kahrp* B) of the original 282bp *kahrp* EMSA probe were designed using SerialCloner (Version 2.6.1, *SerialBasics*), melting temperatures and self-complementarity of the primers was checked using OligoCalc (www.basic.northwestern.edu/biotools/oligocalc.html).

The double stranded copolymer poly(dG-dC) (*Sigma-Aldrich*) was used to compete with the target DNA sequence for non-specific binding of the proteins of interest.

Table 2. DNA probes used in ITAs and EMSAs. PCR primer sequences in UPPERCASE, GAL4 binding sites uppercase *ITALICS*, TSS indicated in **red**. Known TATA-consensus sequences in TdT and Ad2ML promoters indicated in **black**.

EMSA PROBES			
Promoter region	**Sequence**	**Length (bp)**	**Biotinylated**
kahrp	**5'**CGAGCGGAGACTCTAGAATTCtgatgatatgtttatatttcttgttttaaattattagaataaaaagaaataattattatttcatggaaataatataatattattatttttttttttaattatttagtagttatgttttgtcgtttttctcatttatttattataatttacctatagtatatactatgatgtatatttactctagtatgaagaataaagttaatgtaaaatattactacactacatgcagttttaGGATCCGGAGAGCTCCCAACGCGTT**-3'**	282	5' and 3' OR no
gbp-130	**5'**CGAGCGGAGACTCTAGaattctttgaagtacactcaaaataagttatataccatatgtttttaacatatattatatatatatatatatataatataatataatattataattatttttttaatattattaaattgaacataattattttatatcttactattatttttagaaaatttattatatatacatgcaatcataaataatgttttccctgaacctttttcaatgaaataagttaacacaccattccttttGGATCCGGAGAGCTCCCAACGCGTT**-3'**	282	5' and 3' OR no
Ad2ML	**5'-**AATTGGGCCCGACGTCGCAtgctcctctagctgaagcttgcatgcctgcaggtcggagtactgtcctccgagcggagtactgtcctccgagcggagtactgtcctccgagcggagtactgtcctccgagcggagtactgtcctccgagcggagactctaggggtgttcctgaagggggggc**tataaaag**ggggtgggggcgcgttcgtcctcactctcttccgcatcgctgtctgcgagggccagctgttggggtgagtactccctctcaaaaggatccGGAGAGCTCCCAACGCGTT-**3'**	297	5' and 3' OR no
var	**5'-**atatatgtatttattttcatagaaatgtggtagataatatagatagaaaggtaatatttc**-3'**	60	5' and 3'
kahrp A	**5'**-CGAGCGGAGACTCTAGAATTCtgatgatatgtttatatttcttgttttaaattattagaataaaaagaaataattattatttcatggaaataatataatattattatttttttttttaattatttagtagttatgtt**-3'**	138	no
kahrp C	**5'**-tagtagttatgttttgtcgtttttctcatttatttattataatttacctat agtatatactatgatgtatatttactctagtatgaagaataaagttaatgtaaaatattactacactacatgcagttttaGGATCCGGAGAGCTCCCAACGCGTT**-3'**	157	no
TdT TATA	**5'**-CGAGCGGAGACTCTAGAATTCggtacctatg**tataaaat**ggtgagaggacatcagagccctcattctggagacaccacctgatggcacagacagaggatccg GAGAGCTCCCAACGCGTT**-3'**	126	no
TdT ΔTATA	**5'**-CGAGCGGAGACTCTAGAATTCgagctcggtacctatgggtctgctggtgagaggacatcagagccctcattctggagacaccacctgatggcacagacagaggatccgGAGAGCTCCCAACGCGTT**-3'**	120	no

ITA PROBES			
Promoter region	**Sequence**	**Length (bp)**	**Biotinylated**
kahrp	**5'**AATTGGGCCCGACGTCGCA*TGCTCCTCTA GAGCTTGCATGCCTGCAGGTCGGAGTACT GTCCTCCGAGCGGAGTACTGTCCTCCGAG CGGAGTACTGTCCTCCGAGCGGAGTACTGT CCTCCGAGCGGAGTACTGTCCTCCGAGCG GAGACTCTAGAATTC*tgatgatatgttttatatttcttgtttt aaattattagaataaaaagaaataattattatttc**a**tggaaataa tataatattattatttttttttttaattatttagtagttatgttttgtcgttttttc tcatttatttattataatttacctatagtatatactatgatgtatatttac tctagtatgaagaataaagttaatgtaaaatattactacactaca tgcagttttaGGATCCGGAGAGCTCCCAACGCG TT**-3'**	422	5'
gbp-130	**5'**AATTGGGCCCGACGTCGCA*TGCTCCTCTA GAGCTTGCATGCCTGCAGGTCGGAGACTG TCCTCCGAGCGGAGTACTGTCCTCCGAGC GGAGTACTGTCCTCCGAGCGGAGTACTGTC CTCCGAGCGGAGTACTGTCCTCCGAGCGG AGACTCTAGAATTC*tttgaagtacactcaaaataagtta tataccatatgtttttaacatatattatatatatatat**a**tatatataata taatataatattataattatttttttaatattattaaattgaacataatta ttttatatcttactattatttttagaaaatttattatatatacatgcaatc ataaataatgttttccctgaacctttttttcaatgaaataagttaaca caccattccttttGGATCCGGAGAGCTCCCAACGC GTT**-3'**	422	5'
Ad2ML	**5'**AATTGGGCCCGACGTCGCAtgctcctctagctga agcttgcatgcctgcaggtcggagtactgtcctccgagcggagt actgtcctccgagcggagtactgtcctccgagcggagtactgtc ctccgagcggagtactgtcctccgagcggagactctaggggtg ttcctgaaggggggc**tataaaag**ggggtgggggcgcgttcgt cctcactctcttccgcatcgctgtctgcgagggccagctgttggg gtgagtactccctctcaaaaggatccGGAGAGCTCCCA ACGCGTT-**3'**	297	5'

DNA immobilisation on streptavidin magnetic beads

Biotinylated probes were bound to Dynabeads® M-280 streptavidin beads (Dynabeads, *Invitrogen*) following the manufacturer's instructions. Probe-bound beads were resuspended in 1x wash buffer (10mM Tris-Cl pH 7.5 at RT; 0.5mM EDTA; 1M NaCl) containing 100ng/µl bovine serum albumin (BSA). The binding efficiency of the beads was tested by resuspending an aliquot of the beads in deionised water and heating at 70°C for 2min with agitation. Supernatant was separated from the beads using a magnetic separator and resolved by 1.5% agarose gel electrophoresis and quantified using DNA MW marker.

Native PAGE EMSA

Binding reactions of 10µl were prepared combining 4µl template mix prepared in transcription buffer (containing required amount of DNA probe and competitor DNA or poly(dG-dC)) and 6µl protein mix (containing tested proteins in BC-100) for a final transcription buffer composition of 20mM 4-(2- hydroxyethyl)-1-piperazineethanesulfonic acid pH 8.0 (HEPES), 5mM $MgCl_2$, 5mM DTT, 100ng/µl BSA. Binding reactions were incubated at 30˚C, RT or on ice for 45min. Prior to sample loading, 5% 37:1 polyacrylamide gels (0.5xTBE, 0.5mM DTT, 2mM $MgCl_2$) were pre-run at 4˚C for 1h at 100V in 0.5x TBE containing 2mM $MgCl_2$. Complexes were separated by electrophoresis as above at 4˚C for 2h at 100V. Bio-Rad Mini Trans-Blot® Electrophoretic Transfer Cell was used for the transfer of gel to Immobilon-Ny+ nylon membrane (*Merck Millipore*) in ice-cold 0.5x TBE at 4˚C for 2h at 100V. The DNA was crosslinked to the membrane using 4x 120mJ/cm^2 commercial UV-light cross linker and visualised on X-ray film using the Chemiluminescent Nucleic Acid Detection Module (*Thermo Scientific*).

Mg^{2+} agarose EMSA

Binding reactions were prepared as above. Where applicable, antibodies were included in the protein mix and pre-incubated for 15min on ice. Binding reactions were resolved by 1.4% agarose gel (0.5x TBE; 5mM $MgCl_2$) at 4˚C for 2h at 100V in 0.5x TBE containing 2mM $MgCl_2$. The agarose gels were not pre-run prior to sample loading.

A 240mbar diaphragm pump (Charles Austen DA7C) was utilised in conjunction with a southern blotting apparatus (ATTO, USA) to construct a horizontal vacuum-transfer module for Mg^{2+} agarose EMSA gels. The gel was placed on a vacuum-proof mask covering the nylon membrane and subjected to vacuum for 1.5h at RT. Vacuum pressure was controlled by an adapted switch with an adjustable airlock and set to a strength at which from visual observation <50% of the gel was contracted. Following transfer, DNA was crosslinked and visualised as above.

Immobilised template assay

Binding reactions of 40µl were prepared in transcription buffer detailed above using probes

immobilised on streptavidin magnetic beads. Because the protein of interest interacting with DNA is identified by a specific antibody, the amount of protein bound to the DNA has to be sufficient for antibody detection. For this reason, 1pmol of DNA probe immobilised on streptavidin beads was used in this assay. Equivalent amount of free streptavidin magnetic beads were used to control for non-specific protein interactions with the magnetic beads.

The DNA-bound and free beads were equilibrated in standard transcription buffer, followed by incubation with the binding reactions for 45 minutes at 30°C. Following incubation, the unbound fraction was separated from the beads using a magnetic separator and mixed with 15µl 4x Laemmli Sample Buffer (*Bio-Rad*) supplemented with 1M ß-EtSH. The beads were washed in 100µl of transcription buffer and resuspended in 40µl of transcription buffer and 15µl 4x Laemmli Sample Buffer (*Bio-Rad*) supplemented with 1M ß-EtSH. The samples were heated at 80°C for 2 minutes and resolved by SDS PAGE and visualised by immunoblotting, as described above.

3. Results

3.1. Expression and purification of PfTBP

3.1.1. Cloning of 6His:PfTBPc

The expression vector for PfTBP was cloned previously and generously provided by Dr. Thomas Oelgeschläger. This vector expresses 6His- tagged PfTBP amino acids 100-327 (PfTBP 'core' = PfTBPc), containing the evolutionarily conserved C-terminal DNA-binding domain and residues 100-147 of the N-terminus. For a detailed plasmid map of 6His:PfTBPc see Appendix A.

3.1.2. Expression and purification of 6His:PfTBPc

The chosen expression system for the bacterial expression of all recombinant *P. falciparum* proteins was the T7 promoter system expressed in *E. coli* BL21-CodonPlus-(DE3)-RIL cells. The protein of interest is cloned into plasmid pET11d, downstream of the T7 promoter. The BL21-CodonPlus-(DE3)-RIL strain encodes T7 polymerase that is expressed under the control of the IPTG-inducible *lac*UV5 promoter.

Recombinant 6His:PfTBPc expression

6His:PfTBPc expression conditions were established previously in the lab by Robert Milton. Briefly, initial cell mass was generated by an overnight starter culture with media containing 1% (w/v) glucose. Expression media lacking glucose was inoculated with starter culture at a ratio 1:100 (v/v) and incubated at 37˚C until mid-log phase (OD_{550} = ~0.5). The cultures were cold-shocked on ice for 90s and incubated at 30˚C in the absence of IPTG inducer for 16 hours. The cultures were not induced with IPTG, as similarly to the PfTFIIB expression (Chapter 3.4), IPTG induction led to reduced growth rate and cell death (Robert Milton, unpublished observations). A detailed expression protocol is provided in Chapter 2.2.

Purification of 6His:PfTBPc

6His:PfTBPc was affinity purified via the 6His-tag using Ni-NTA agarose resins. The Ni-NTA agarose elution profile of 6His:PfTBPc purified from cleared bacterial lysate was broad and preliminary immobilised template assays showed that the late fractions had significantly higher

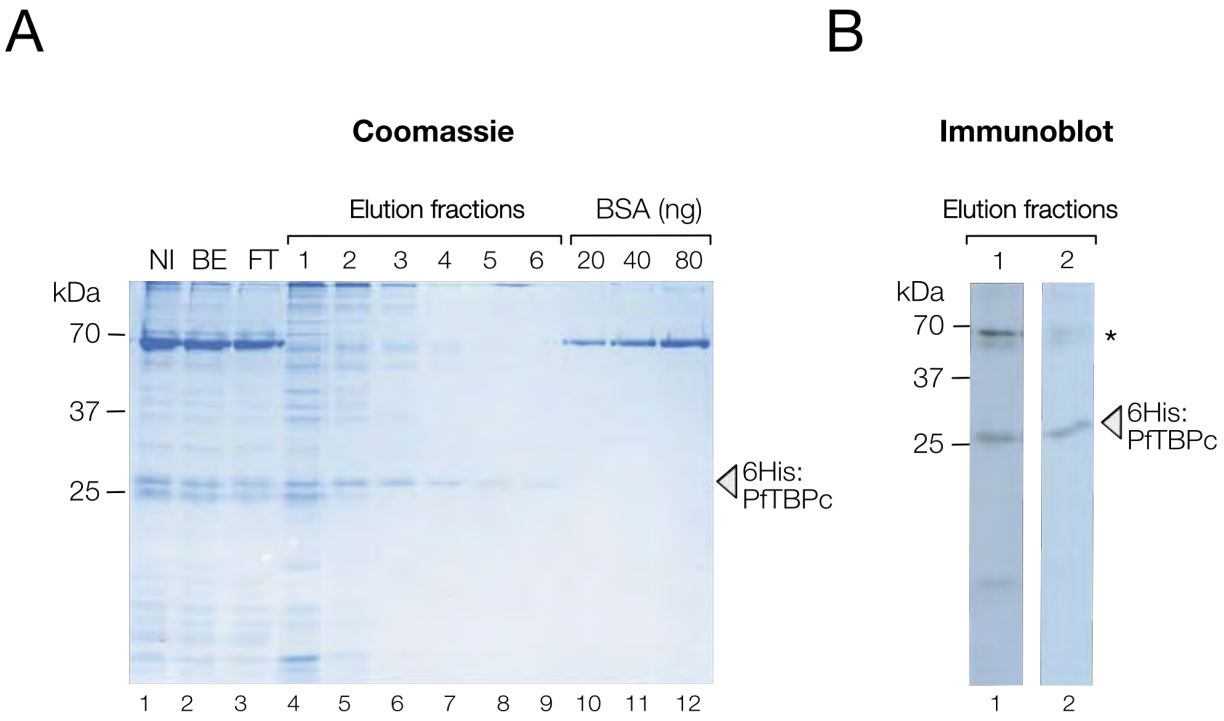

Figure 6. 6His:PfTBPc purification with Nickel affinity resins

A. SDS PAGE analysis of PfTBPc fractions, purified by Ni-NTA resin and nickel magnetic beads.

6His:PfTBPc was purified by Ni-NTA agarose from cleared bacterial lysates, buffer exchanged into BC-500 (500mM KCl), followed by purification with Ni-NTA magnetic beads. Eluates were analysed by SDS PAGE with comparison to known amounts of BSA. Equivalent amounts of pooled Ni-NTA agarose eluate before (NI) and after buffer exchange to 500mM KCl (BE), Ni-NTA magnetic bead flowthrough (FT) and 10µl (1/80th of the total elution volume) of each elution fraction was resolved by 15% SDS PAGE. The relative position of protein MW standards (BioRad Precision Plus All Blue) and the band corresponding to 6His:PfTBPc is indicated.

B. Immunoblot analysis of 6His:PfTBPc purification on Ni-NTA beads.

Ni-NTA magnetic bead purification elution fractions 1 and 2 (3µl; corresponding to lanes 4,5; panel A) were resolved by 15% SDS PAGE, transferred to a PVDF membrane and probed with a 1:500 dilution of α-PfTBP rabbit antisera. X-ray film exposure time: 10s. The 67kDa band indicated by an asterisk (*), corresponds to BSA added to the fraction at a concentration of 100ng/µl, which at this high concentration cross-reacts with the α-PfTBP rabbit antisera.

DNA-binding activity compared to the early fractions (data not shown). It is possible that the reduced binding activity was caused by aggregation due to high PfTBPc concentration in the early fractions. In order to reduce ionic interactions between proteins in solution and alleviate aggregation, the Ni-NTA eluates were pooled, adjusted to 500mM KCl using desalting column (PD-10, *GE Healthcare*), and re-purified using Ni-NTA magnetic beads. 6His:PfTBPc eluted from Ni-NTA magnetic beads mostly in the first fractions (Fig. 6A). Selected fractions were analysed with α-PfTBP rabbit antisera (Fig. 6B), confirming the identity of 6His:PfTBPc.

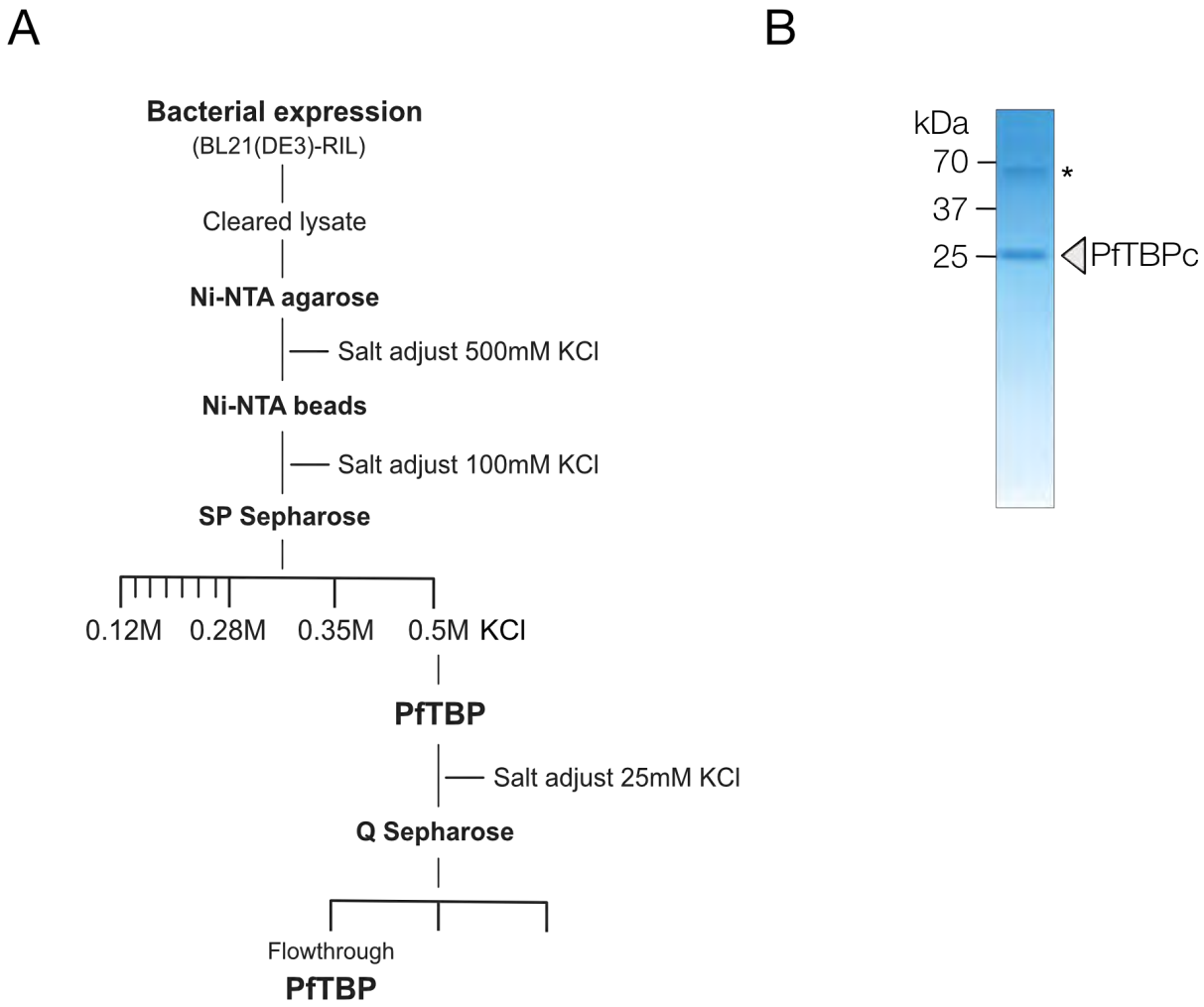

Figure 7. PfTBPc purification using ion chromatography.
A. Overall 6His:PfTBPc purification scheme
B. SDS PAGE analysis of 6His:PfTBPc, purified according to the presented scheme (details provided in Chapter 2.3). Q-Sepharose flowthrough (15µl) was resolved by 5%/15% SDS PAGE. A contaminant at 70kDa indicated by an asterisk (*). The relative position of protein MW standards (BioRad Precision Plus All Blue) is indicated.

The purity of the 6His:PfTBPc preparation was not sufficient at this stage for electrophoretic mobility shift assays and further purification with ion exchange resins was carried out by Dr. Thomas Oelgeschläger. PfTBPc has a calculated pI of 9.3 and is therefore expected to bind to cation exchange resins. Briefly, the Ni-NTA elutions were pooled, adjusted to 100mM KCl using desalting column (PD-10, *GE Healthcare*), bound to strong cation exchange resin, SP-Sepharose Fast Flow (*Sigma-Aldrich*), and eluted with 500mM KCl. The eluates were adjusted to 20mM KCl as above, and passed over Q-Sepharose Fast Flow (*Sigma-Aldrich*), in order to bind remaining contaminants and recover purified 6His:PfTBPc in the flowthrough (purification scheme presented in Fig. 7A). As seen in Fig. 7B, the majority of the Ni-NTA co-purified contaminants were removed by ion chromatography and the 6His:PfTBPc preparation contains only one detectable protein contaminant at ~70kDa. The 6His:PfTBPc preparation presented in Fig. 7B was utilised in subsequent DNA binding assays.

3.2. Initial characterisation of PfTBP in DNA binding assays

3.2.1. PfTBP binds to immobilised *gbp-130* and Ad2ML promoter DNA

Initial assessment of PfTBP binding activity was performed by immobilised template assay (ITA), that allows direct detection of DNA-bound protein with specific antibodies by immunoblot. 6His:PfTBPc was incubated with a putative *P. falciparum* promoter *gbp-130* or with the eukaryotic model promoter, adenovirus 2 major late (Ad2ML), immobilised on magnetic beads. To control for non-specific interactions of PfTBP with the magnetic beads, PfTBP was also incubated with free beads. After separating the bound and unbound fractions, the bound fractions were washed extensively in order to eliminate any residual unbound proteins and analysed using α-PfTBP antibody. As seen from Figure 8, the majority of PfTBP was detected in the bound fractions for both *gbp-130* and Ad2ML DNA templates. Conversely, the majority of PfTBP did not bind to the free beads, showing that the protein detected on *gbp-130* and Ad2ML was interacting with the immobilised target DNA, rather than the beads. These results provide strong evidence that the purified PfTBP was indeed active and had DNA-binding activity. Furthermore, PfTBP bound to *gbp-130* and Ad2ML promoters to a similar extent, suggesting that both DNA probes are bound with similar binding affinity under these conditions.

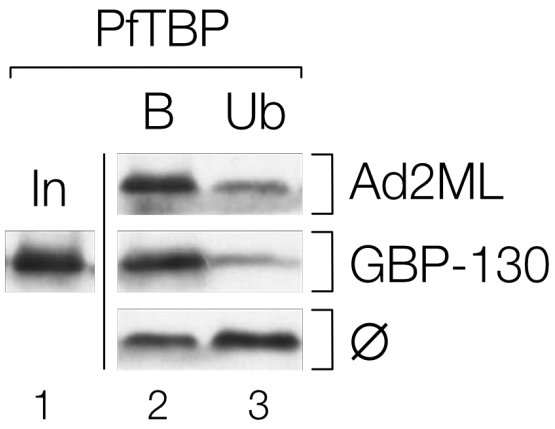

Figure 8. PfTBP binds to Ad2ML and *gbp-130* promoters.
Immunoblot analysis of PfTBP binding to immobilised *gbp-130* and Ad2ML model promoters. DNA-binding reactions (40µl) contained 50ng 6His:PfTBPc and 1pmol of immobilised Ad2ML, *gbp-130* promoter DNA or free beads, as indicated. Input (20ng; **In**), 50% of unbound (**Ub**; lane 3) and an equivalent amount of bound fraction (**B**; lane 2) were analysed by immunoblot with α-PfTBP antibody (1:500). X-ray film exposure: 30s. Note that all panels originate from the same gel with the same input.

To assess the DNA sequence specificity of PfTBP, PfTBP binding to the *gbp-130* promoter was challenged with poly(dG-dC) competitor DNA. Minimally, a 8-fold molar excess of competitor was required to disrupt PfTBP binding to *gbp-130* (data not shown). The high excess of non-specific competitor DNA necessary to outcompete *gbp-130* for PfTBPc binding may imply preferential PfTBP binding to A/T-rich DNA sequences.

Taken together, these results confirm the direct binding of PfTBP to the model promoter sequences, providing the basis for further investigation of DNA-binding by EMSA.

3.2.2. Temperature dependence of PfTBP DNA-binding

Initial EMSAs with *Plasmodium* proteins using native polyacrylamide gels resulted in quantitative probe retention in the wells (data not shown). Protein concentration-dependent decrease of the free probe corresponded to the increase of probe retained in the wells. Therefore, it was concluded that the proteins under investigation possess DNA-binding activity that cannot be resolved under these conditions. Variation of the acrylamide:*bis*-acrylamide ratio, electrophoresis conditions, and KCl concentration of the binding reactions, did not help to

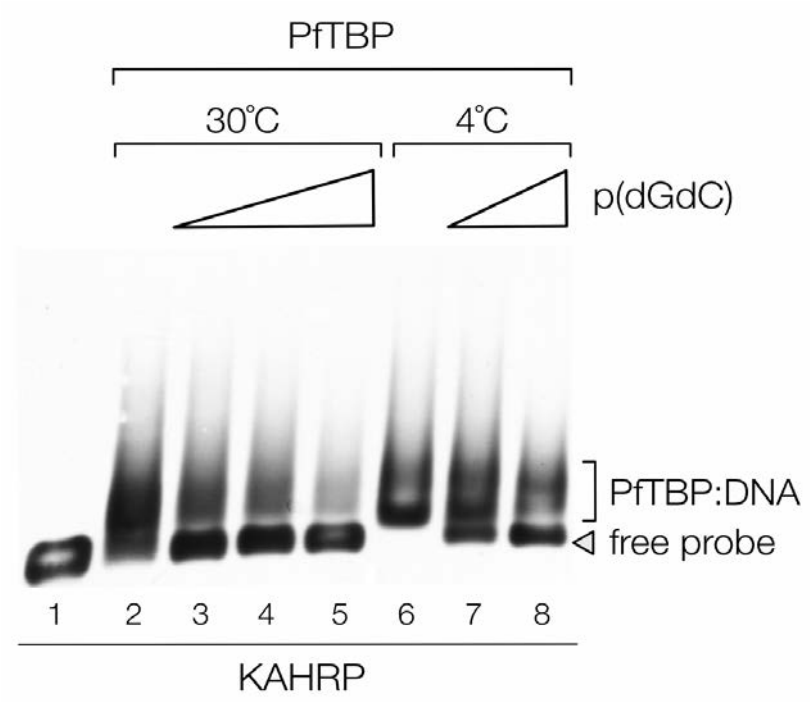

Figure 9. PfTBP-DNA complex formation is temperature-dependent.
Analysis of 6His:PfTBPc binding to the *kahrp* promoter by Mg^{2+}agarose EMSA.
Binding reactions (10µl) contained 20fmol 282bp *kahrp* promoter DNA, 100ng (lane 3), 200ng (lanes 4, 7) or 400ng (lanes 5, 8) poly(dG-dC), as indicated, and 12ng of 6His:PfTBP. Binding reactions were incubated at 30˚C or 4˚C for 45min, as indicated. Protein-DNA complexes were resolved in 1.4% agarose gels (0.5xTBE; 5mM $MgCl_2$). X-ray film exposure: 2s.

resolve protein-bound probe in the gel. For this reason, the protocol was adjusted for Mg^{2+} agarose gels (for details see Chapter 2.5), that is used as an alternative to non-denaturing polyacrylamide gels (Hellman and Fried, 2007), as larger pore size of agarose allows for the resolution of large multi-protein complexes (Zerby and Lieberman, 1997).

In order to assess PfTBP DNA-binding activity in Mg^{2+} agarose EMSA, 10fmol of *kahrp* promoter DNA was incubated with 12ng of 6His:PfTBPc in standard transcription conditions. The protein-DNA complex formation was highly variable between different experiments (data not shown). The variable duration of binding reactions on ice prior to electrophoresis correlated with the differential protein-DNA complex formation. Therefore, the effect of the incubation of 6His:PfTBPc with DNA probe at standard 30˚C and on ice (~4˚C) was compared. As seen in Figure 9, binding of PfTBP and smearing of the probe was observed after 30˚C incubation. In order to assess the stability of the complexes, poly(dG-dC) competitor DNA was included in the binding reactions. In the presence of 400ng poly(dG-dC), complete disruption of the complex

formed at 30˚C was observed, with the majority of the free probe recovered (Fig. 9, compare lanes 2 and 5). This is consistent with the findings in ITAs using similar amount of competitor. Upon incubation on ice, DNA probe was quantitatively bound by PfTBP, forming a clear gel shift with significantly reduced smearing compared to the complex formed at 30˚C. Challenge with 400ng poly(dG-dC) competitor DNA recovered some free probe, with an estimated 50% of probe remained in the complex (Fig. 9, compare lanes 6 and 8).

The difference in appearance of the mobility shifts formed at 30˚C and on ice, and sensitivity to poly(dG-dC) competitor indicates a quantitative and qualitative difference in PfTBP binding. Firstly, PfTBP binds more probe when binding reactions are incubated on ice (Fig. 9, compare lanes 2 and 6). It seems possible, that 6His:PfTBPc incubation at higher temperatures might cause partial denaturation, leading to reduced DNA-binding activity. Secondly, complexes formed on ice have a lesser extent of probe smearing and are less sensitive to challenge by poly(dG-dC) competitor compared to complexes formed at 30˚C (Fig. 9, compare lanes 3-5 and lanes 7-8). Smearing is observed in electrophoresis conditions with unstable complexes, where rapid dissociation of the protein from the DNA does not allow for clear resolution of the mobility shift. Therefore, besides increased binding of the probe, PfTBP complexes formed on ice also appear to have higher stability than complexes formed at 30˚C. Taken together, in contrast to previously characterised eukaryotic TBPs that exhibit optimal binding at 30˚C (Ge et al., 1996), our findings indicate a preference of PfTBP for lower incubation temperatures, with higher incubation temperatures most likely affecting the intrinsic properties of the protein.

3.2.3. PfTBP binds *kahrp* and Ad2ML promoter regions with similar affinity

In order to further investigate the PfTBP-DNA complexes, Ad2ML or *kahrp* promoters were incubated with increasing amounts of PfTBP in the absence of competitor DNA. With 9ng of PfTBP both promoters were quantitatively bound and the PfTBP-promoter complexes were seen migrating directly above the free probe (Fig. 10). No significant difference in quantity or mobility was observed between the PfTBP nucleoprotein complexes with *kahrp* and Ad2ML promoters. As seen on the lower exposure (Fig. 10, 1s expo), increase of PfTBP resulted in diminishing of the probe, indicative of low-affinity binding across the target DNA. Interestingly,

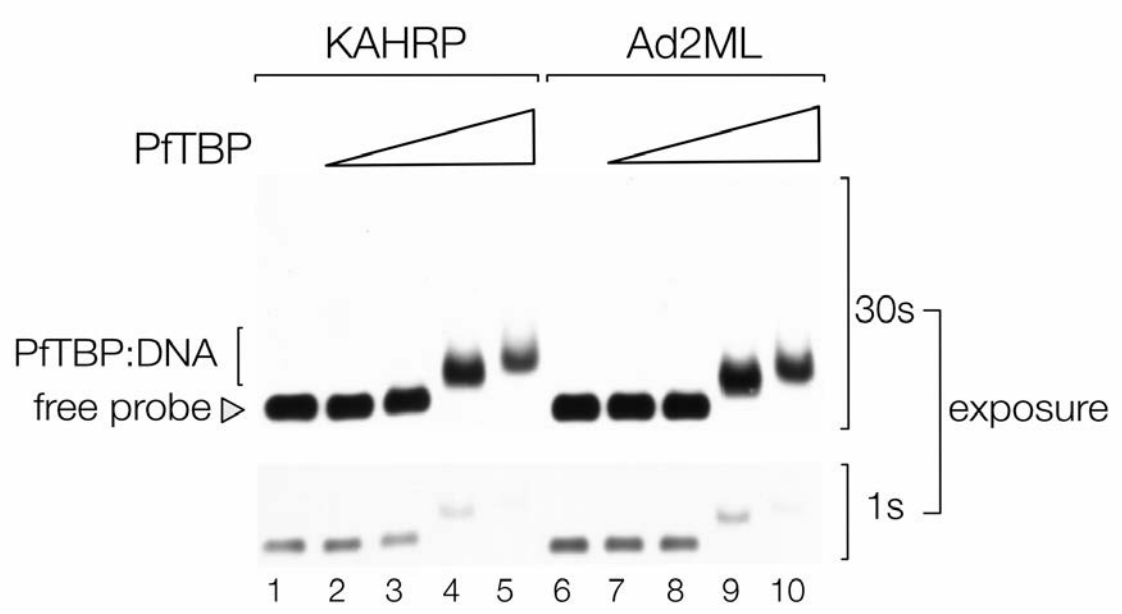

Figure 10. PfTBP has similar binding activity on *kahrp* and Ad2ML promoters.
Analysis of PfTBP binding to *kahrp* and Ad2ML promoters by Mg^{2+}agarose EMSA. Binding reactions (10µl) contained 5fmol 282bp *kahrp* or 297bp Ad2ML probe and 0.9ng (lanes 2,7), 2.7ng (lanes 3,8), 9ng (lanes 4,9) or 27ng (lanes 5,10) of 6His:PfTBPc. Protein-DNA complexes were resolved in 1.4% agarose gels (0.5xTBE; 5mM $MgCl_2$). The two panels indicate 1s and 30s exposure of the same experiment.

---

the relative mobility of the PfTBP-DNA decreased with increasing amounts of PfTBP. This could indicate successive binding of PfTBP to multiple sites, determined by PfTBP concentration.

3.3. Expression and purification of 6His:PfTFIIB:GST

3.3.1. Cloning of 6His:PfTFIIB:GST

PfTFIIB has been cloned previously within our research group and a potential C-terminal degradation upon recombinant expression was observed. To ensure the purification of full-length PfTFIIB, an expression vector for N-terminally 6His-tagged and C-terminally GST-tagged full-length PfTFIIB was constructed. This was achieved by construction of an intermediate vector by insertion of a GST ORF into a pET11d-6His plasmid, and subsequent ligation of PfTFIIB ORF into the constructed pET11d-6His-GST vector frame.

The open reading frame for PfTFIIB was obtained by PCR amplification of cDNA phage library MRA-296. The MRA-296 cDNA phage library is derived from mRNA of blood stage cultures of *Plasmodium falciparum* clone 3D7 16 hours post-synchronisation, corresponding to the late-ring stage of the parasite (Chakrabarti et al., 1994).

The detailed cloning protocol is provided in Chapter 2.1 and the plasmid map for pET11d-6His-PfTFIIB-GST can be found in Appendix A.

3.3.2. IPTG-induced expression of 6His:PfTFIIB:GST inhibits cell growth

Previous studies within our research group have linked IPTG-induction of the heterologous expression of *P. falciparum* proteins to the inhibition of host cell growth. In order to assess the effect of IPTG-induced expression of 6His:PfTFIIB:GST on the host cells, two independent expression cultures were set up. *E. coli* BL21-codonPlus (DE3)-RIL cells were transformed with

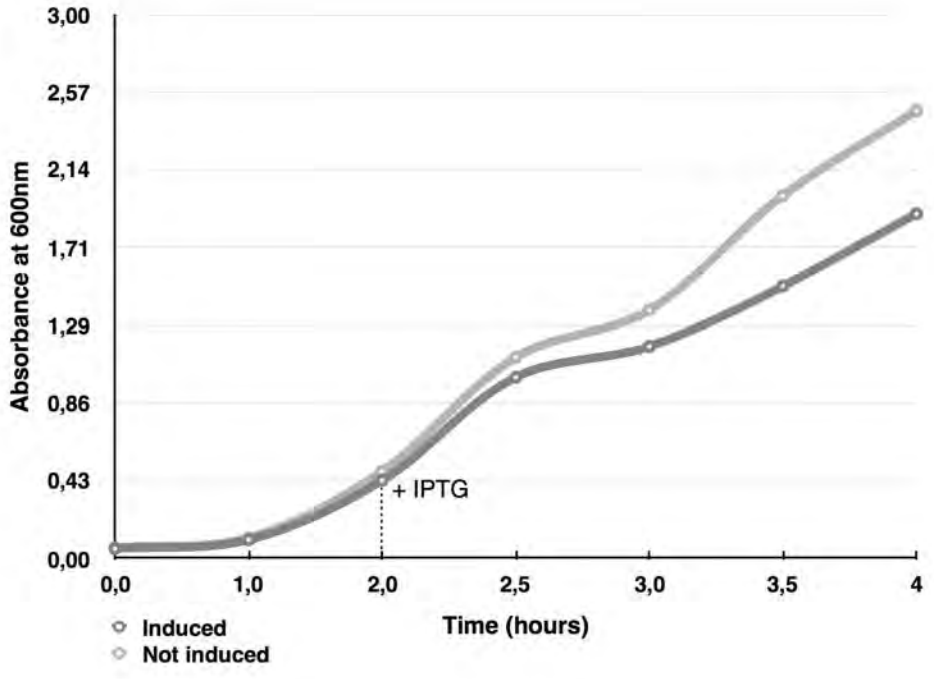

Figure 11. IPTG-induced 6His:PfTFIIB:GST expression in *E.coli* BL21(DE3)-RIL cells leads to growth inhibition.
Cell densities were measured spectrophotometrically at a wavelength of 600nm and plotted against time. Expression cultures were induced with 1mM IPTG after 2 hours of growth. Growth of the induced (dark grey) and uninduced (light grey) cultures was compared.

pET11d-6His-PfTFIIB-GST expression vector and grown in liquid media at 37°C. The cell density was measured spectrophotometrically throughout initial growth phases and, at log-phase (OD_{600nm}= ~ 0.5), one of the two cultures was induced with 1mM IPTG. Cell growth was continuously monitored post-induction and cells harvested at stationary phase. IPTG-induced cultures exhibited a reduction in growth, observable immediately post-induction with IPTG (Fig. 11). The detrimental effect to cell growth could be due to PfTFIIB toxicity to the cells and thus the expression levels of 6His:PfTFIIB:GST in induced and uninduced cultures were investigated.

3.3.3. Optimal 6His:PfTFIIB:GST expression is achieved from uninduced cultures

In order to assess the expression of 6His:PfTFIIB:GST in induced and uninduced cultures, the expressed proteins were purified via the 6His-tag using Ni-NTA magnetic beads or via the GST-tag using GSH-agarose, as described in Chapter 2.2. A 70kDa band, which is the expected size of 6His:PfTFIIB:GST, was enriched with both affinity purification methods, providing strong evidence for recombinant 6His:PfTFIIB:GST expression (Fig. 12). Notably, 6His:PfTFIIB:GST

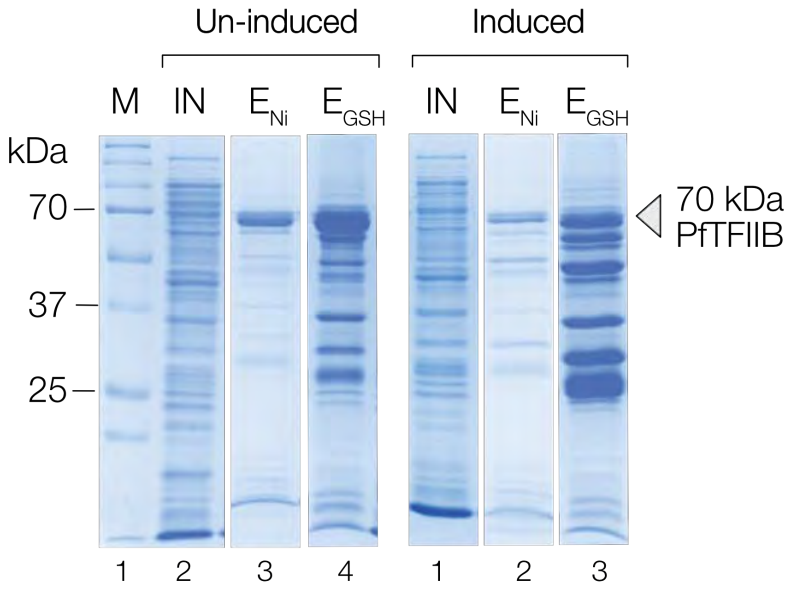

Figure 12. 6His:PfTFIIB:GST is expressed in induced and uninduced cultures.
SDS PAGE analysis of 6His:PfTFIIB:GST expression in induced and uninduced cultures. Cleared lysate (input, **IN**) of induced and uninduced cultures was passed over Ni-NTA beads or GSH-beads. Eluates (**E_{NI}** and **E_{GSH}**, respectively; 10µl [=1/5th total elution volume]) were resolved by 12% SDS PAGE. **M**: Molecular weight marker (BioRad Precision Plus All Blue). The relative position of protein MW standards is indicated.

was effectively expressed in uninduced cultures, where also more cell mass could be harvested. It seems, in the absence of IPTG, the basal level of T7 polymerase expression induces PfTFIIB expression at levels that are not toxic to the host cell and thus enables the generation of more cell mass compared to the IPTG-induced cultures. These observations are consistent with previous findings within our research group (Steven Bing, unpublished observations). As sufficient expression and higher cell mass was obtained from uninduced cultures, it was deemed more efficient and established as a standard for PfTFIIB expression.

3.3.4. Sequential affinity-purification of 6His:PfTFIIB:GST

Further purification of 6His:PfTFIIB:GST was attempted using Ni-NTA and GSH agarose in a sequential manner, as described in Chapter 2.2. Two dominant bands migrating at 70kDa and around 25kDa were seen after the purification with GSH agarose (Fig. 13). The band migrating around 70kDa, the expected size of 6His:PfTFIIB:GST, remained prominent after Ni-NTA

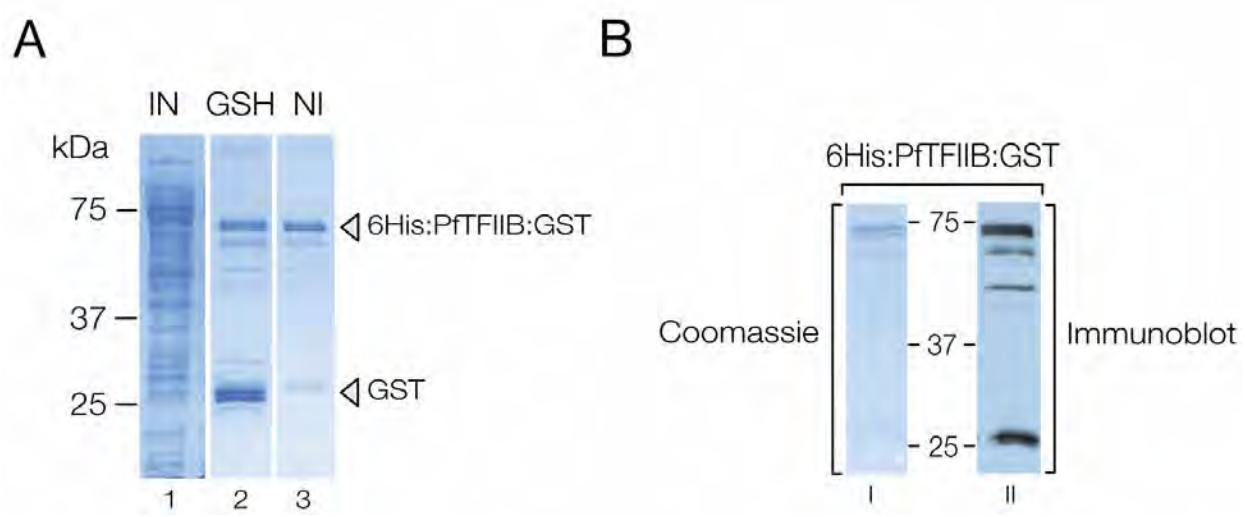

Figure 13. Purification of 6His:PfTFIIB:GST by sequential affinity chromatography.
A. SDS PAGE analysis of purified 6His:PfTFIIB:GST fractions. Cleared lysate (input, **IN**) of uninduced cultures was purified on GSH-agarose, followed by Ni-NTA agarose. Eluates (**GSH** and **NI**, respectively, 10µl [=1/300th total elution volume]) were resolved by 12% SDS PAGE. The relative position of protein MW standards (BioRad Precision Plus All Blue), 6His:PfTFIIB:GST and GST-tag is indicated.
B. Immunoblot analysis of purified 6His:PfTFIIB:GST. 6His:PfTFIIB:GST (40ng) was resolved by 12% SDS PAGE, transferred to a PVDF membrane and probed with a 1:1000 dilution of affinity-purified α-PfTFIIB antibody. X-ray film exposure time: 10s. Coomassie-stained gel that corresponds to the immunoblot is indicated.

purification. GST-tag has a molecular weight of 26kDa, corresponding to the prominent band seen after GSH-agarose purification. The presence of free GST-tag is commonly reported in GST-fusion protein expression systems and could be the result of intracellular proteolytic degradation events that lead to premature excision of the tag from the fusion protein (Harper and Speicher, 2011).

6His:PfTFIIB:GST fusion protein migrated in SDS PAGE gels as a doublet of bands and analysis with purified α-PfTFIIB (identifies residues 353-367) confirmed the immunoreactivity of both bands of the doublet (Fig. 13B). This feature of PfTFIIB is discussed further in Chapter 3.4.3. In addition to 6His:PfTFIIB:GST, the purified α-PfTFIIB antibody detected low molecular weight protein bands in the preparation, indicating the presence of PfTFIIB proteolytic breakdowns.

The 6His:PfTFIIB:GST preparation presented in Figure 13B was utilised in subsequent DNA binding assays.

3.4. Expression and purification of 6His:PfTFIIB, 6His:PfTFIIBΔ37 & 6His:PfTFIIBΔ140

Initial investigation into PfTBP-PfTFIIB complex formation in EMSA showed PfTBP-independent DNA-binding activity of PfTFIIB. In order to determine the regions of PfTFIIB harbouring DNA-binding activity, two PfTFIIB deletion mutants were generated. PfTFIIB is predicted to harbour a highly basic (pI>11.5) 37-residue N-terminal region. Structural predictions indicate this region to be a largely unstructured non-conserved region, with some potential for α-helix formation (Steven Bing, unpublished observations). This region could contribute to non-specific DNA binding by interacting with the negatively charged DNA backbone, however, specific interactions via alternative DNA-binding motifs cannot be excluded. The C-terminal domain is predicted to contain two cyclin-like repeats (Nikolov et al., 1995) which in other eukaryotes harbour a helix-turn-helix motif, that is involved in DNA sequence recognition (Lagrange et al., 1998). With these features in mind, deletion mutants lacking the N-terminal highly basic 37 residues (termed PfTFIIBΔ37) or the entire N-terminus (termed PfTFIIBΔ140) were generated.

3.4.1. Cloning of 6His:PfTFIIB, 6His:PfTFIIBΔ37 & 6His:PfTFIIBΔ140

The open reading frames of full length PfTFIIB and deletion mutants lacking 37 and 140 and N-terminal amino acids (PfTFIIBΔ37 and PfTFIIBΔ140, respectively), were amplified by PCR from the previously constructed 6His:PfTFIIB:GST vector and ligated into pET11d-6His vector. Primer sequences and detailed cloning protocol are provided in Chapter 2.1. Detailed pET11d-6His-PfTFIIB, pET11d-6His-PfTFIIBΔ37 and pET11d-6His-PfTFIIBΔ140 plasmid maps can be found in Appendix A.

3.4.2. Expression of 6His:PfTFIIB, 6His:PfTFIIBΔ37 & 6His:PfTFIIBΔ140

As discussed in Chapter 3.3.3, 6His:PfTFIIB:GST expression was achieved by auto-induction from glucose-containing media. However, highly variable results were obtained with 1%(w/v) glucose in expression media of 6His-tagged PfTFIIB variants (data not shown). These variable results could be explained by variations in glucose concentration and/or quality, affecting the dynamics of cellular metabolism or inducing a lag in the uptake of glucose from the media, resulting in aberrant expression of PfTFIIB. For these reasons, glucose was used for culturing plates and starter culture media, however it was omitted from expression culture media. Briefly, initial cell mass was generated by an overnight starter culture with media containing 1%(w/v) glucose. Expression media lacking glucose was inoculated with starter culture (ratio 1:20 (v/v)) and incubated at 37˚C for 5 hours. Expression cultures were not induced with IPTG. A detailed expression protocol is provided in Chapter 2.2.

3.4.3. Nickel affinity purification of 6His:PfTFIIB, 6His:PfTFIIBΔ37 & 6His:PfTFIIBΔ140

Full length 6His:PfTFIIB and the mutants 6His:PfTFIIBΔ37 and 6His:PfTFIIBΔ140 were purified via the 6His-tag using Ni-NTA agarose resins. Consistent with the results obtained after the purification of 6His:PfTFIIB:GST, a doublet band at 44kDa separated by approximately 2kDa was observed in the 6His:PfTFIIB elution (Fig. 14A). Immunoblot analysis with purified α-PfTFIIB antibody, that identifies the C-terminal region (residues 353-367) of PfTFIIB, confirmed the doublet bands as PfTFIIB. Upon purification of the deletion mutant 6His:PfTFIIBΔ37, a similar doublet was observed migrating around 40-42kDa (Fig. 14B). As

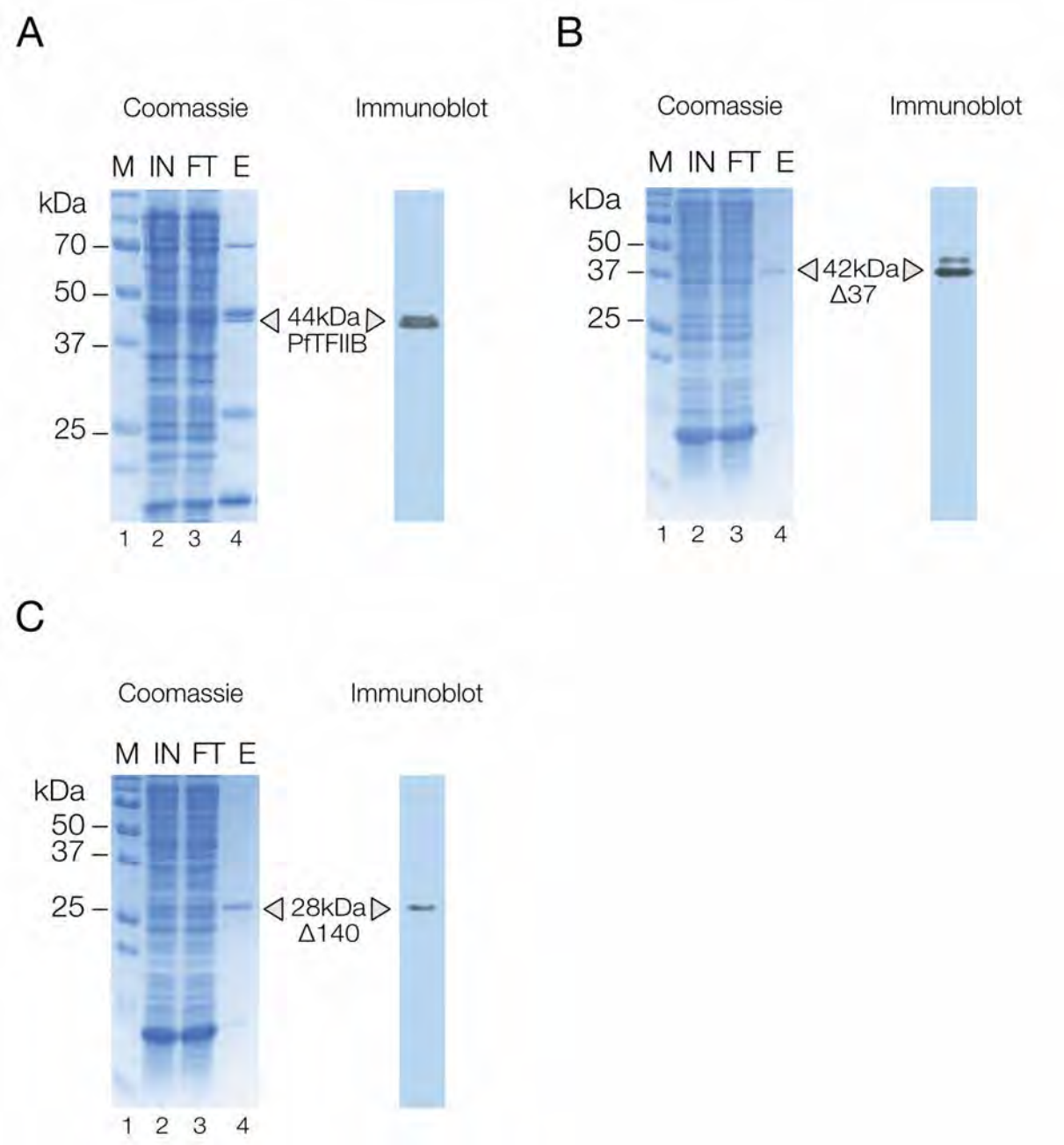

Figure 14. Nickel-affinity purification of full length PfTFIIB and PfTFIIB deletion mutants.

SDS PAGE and immunoblot analysis of Ni-NTA agarose purified 6His:PfTFIIB (**A**), 6His:PfTFIIBΔ37 (**B**) 6His:PfTFIIBΔ140 (**C**).

Equivalent amounts of cleared lysate input (**IN**), Ni-NTA flowthrough (**FT**) and 10µl (1/60th of total elution volume) of Ni-NTA agarose elution (**E**) was resolved by 12% SDS PAGE.

6His:PfTFIIB (5ng), 6His:PfTFIIBΔ140 (5ng) and 6His:PfTFIIBΔ37 (2ng) were resolved by 12% SDS PAGE, transferred to a PVDF membrane and probed with a 1:1000 dilution of affinity-purified α-PfTFIIB antibody. X-ray film exposure time: 10s. **M**: Molecular weight marker (BioRad Precision Plus All Blue). The relative position of protein MW standards and purified recombinant proteins are indicated.

seen in Figure 14C, a single 28kDa band was enriched following Ni-NTA purification of 6His:PfTFIIBΔ140, which lacks the entire N-terminus of PfTFIIB. The presence of the doublet band in all PfTFIIB variants except the PfTFIIBΔ140, suggests that the banding pattern observed could be determined by the zinc-finger structure.

3.4.4. Further purification of 6His:PfTFIIB variants using ion exchange chromatography

In order to achieve sufficient purity for the usage of PfTFIIB variants in electrophoretic mobility shift assays, further purification of the 6His:PfTFIIB was attempted. 6His:PfTFIIB, 6His:PfTFIIBΔ37 and 6His:PfTFIIBΔ140, with calculated pIs of 9.82, 9.37 and 9.82, respectively, are expected to bind to strong cation exchange resins. Furthermore, *E.coli* proteins

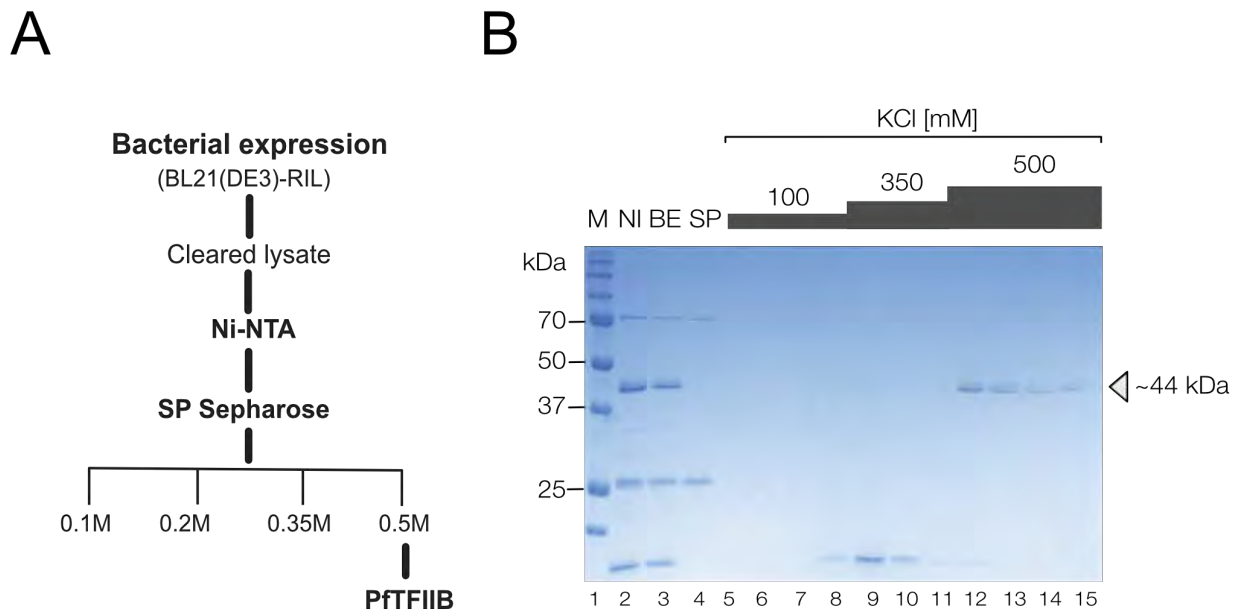

Figure 15. Ion exchange chromatography purification of full length PfTFIIB.
Overall purification scheme of PfTFIIB variants (**A**) and SDS PAGE analysis of full length PfTFIIB purification with SP-Sepharose (**B**).
Cleared lysate was purified by Ni-NTA agarose, eluates adjusted to BC-100 (100mM KCl), and subjected to purification with SP-Sepharose. Equivalent amount of pooled Ni-NTA agarose eluate before (NI) and after buffer exchange to 100mM KCl (BE), SP-Sepharose flowthrough (SP) and 10µl (1/200th of the total elution volume) of the elution fractions (lanes 6-8: 100mM KCl; 9-11: 350mM KCl; 12-15: 500mM KCl) were resolved by 12% SDS PAGE. M: Molecular weight marker (BioRad Precision Plus All Blue). The relative position of protein MW standards and purified recombinant proteins are indicated.

that commonly contaminate Ni-NTA purified proteins, have a pI below 7 (Bolanos-Garcia and Davies, 2006). Pilot purification was attempted using SP-Sepharose Fast Flow (*Sigma-Aldrich*). 6His:PfTFIIB was bound to the resin at 100mM KCl and eluted stepwise with increasing KCl concentrations (Fig. 15A). Two prominent contaminants present in the 6His:PfTFIIB preparation migrating at around 70kDa and 25kDa did not bind the cation exchange resin and were removed in the flowthrough (Fig. 15B, lane 4). The low molecular weight band migrating at around 15kDa is likely to correspond to lysozyme. With a pI of ~11, lysozyme is expected to bind to SP-Sepharose under the conditions. As seen in Figure 15B, 6His:PfTFIIB bound quantitatively to SP Sepharose under the conditions. With stepwise elution using increasing KCl concentrations, lysozyme was removed at 350mM KCl, followed by 6His:PfTFIIB elution at 500mM KCl. All contaminants visible by Coomassie staining were effectively separated from 6His:PfTFIIB using this purification scheme and similar purity was achieved for 6His:PfTFIIBΔ37 and 6His:PfTFIIBΔ140. The purified 6His:PfTFIIB, 6His:PfTFIIBΔ37 and 6His:PfTFIIBΔ140

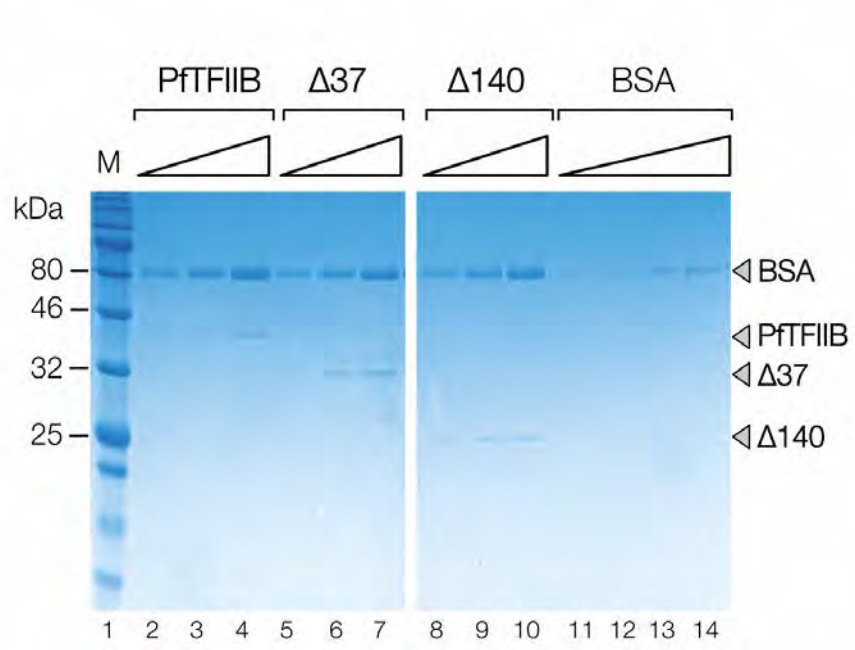

Figure 16. Purified PfTFIIB variants used in EMSAs.
SDS PAGE analysis of purified 6His:PfTFIIB, 6His:PfTFIIBΔ37 and 6His:PfTFIIBΔ140.
SDS PAGE analysis of full length 6His:PfTFIIB (PfTFIIB), 6His:PfTFIIBΔ37 (Δ37) and 6His:PfTFIIBΔ140 (Δ140), purified by Ni-NTA affinity chromatography, followed by ion exchange chromatography on SP-Sepharose. Pooled 500mM KCl SP-Sepharose elutions (1, 2 and 4µl) and 20, 40, 80, 100ng of BSA (lanes 11-14, respectively) were resolved by 12% SDS PAGE. Note that BSA was added to all elution fractions at a concentration of 100ng/µl. **M**: Molecular weight marker (NEB Colour Pertained Protein Standard). The relative position of protein MW standards and purified recombinant proteins are indicated.

preparations were quantified by comparison to BSA standards in SDS PAGE (Fig. 16) and used in subsequent DNA binding assays.

3.5. Initial characterisation of PfTFIIB in DNA-binding assays

3.5.1. PfTFIIB binds immobilised *kahrp* promoter in the absence of PfTBP

The DNA-binding potential of full length and mutant PfTFIIBs was first assessed in immobilised template assays (ITAs). For that purpose, 40ng of each PfTFIIB variant was incubated in standard transcription buffer containing 1pmol of immobilised *kahrp* probe or equivalent amount of free beads. Strikingly, the full length 6His:PfTFIIB as well as the 6His:PfTFIIBΔ37 deletion mutant bound nearly quantitatively to the immobilised DNA under the conditions, whereas no binding was observed to the free beads (Figure 17B, compare lanes 2 and 3). Hence, the deletion of 37 N-terminal residues does not significantly alter the DNA-binding activity of PfTFIIB in this assay. However, only about 50% of the input 6His:PfTFIIBΔ140 bound the DNA probe. These results indicate that the C-terminal region of PfTFIIB possesses PfTBP-independent DNA-binding activity, which is further enhanced minimally by the presence of the N-terminal region lacking the 37-residues.

Challenges with poly(dG-dC) competitor completely disrupted the binding of 6His:PfTFIIB as well as 6His:PfTFIIBΔ140 when 8-fold molar excess of competitor was present (Fig. 17C). The binding specificity of these regions cannot be determined at this stage, as this assay does not discriminate between specific and non-specific binding.

3.5.2. PfTFIIB binds *P. falciparum kahrp* and *var* promoter regions in PAGE EMSA

Next, PfTBP-independent PfTFIIB DNA-binding was investigated in EMSA conditions. In contrast to PfTBP and 6His:PfTFIIB:GST, PfTFIIB-DNA complexes with the 6His-tagged PfTFIIB variants could be resolved in polyacrylamide gel EMSAs. Having previously encountered extreme probe retention using PfTBP and PfTFIIB in conjunction with 282bp

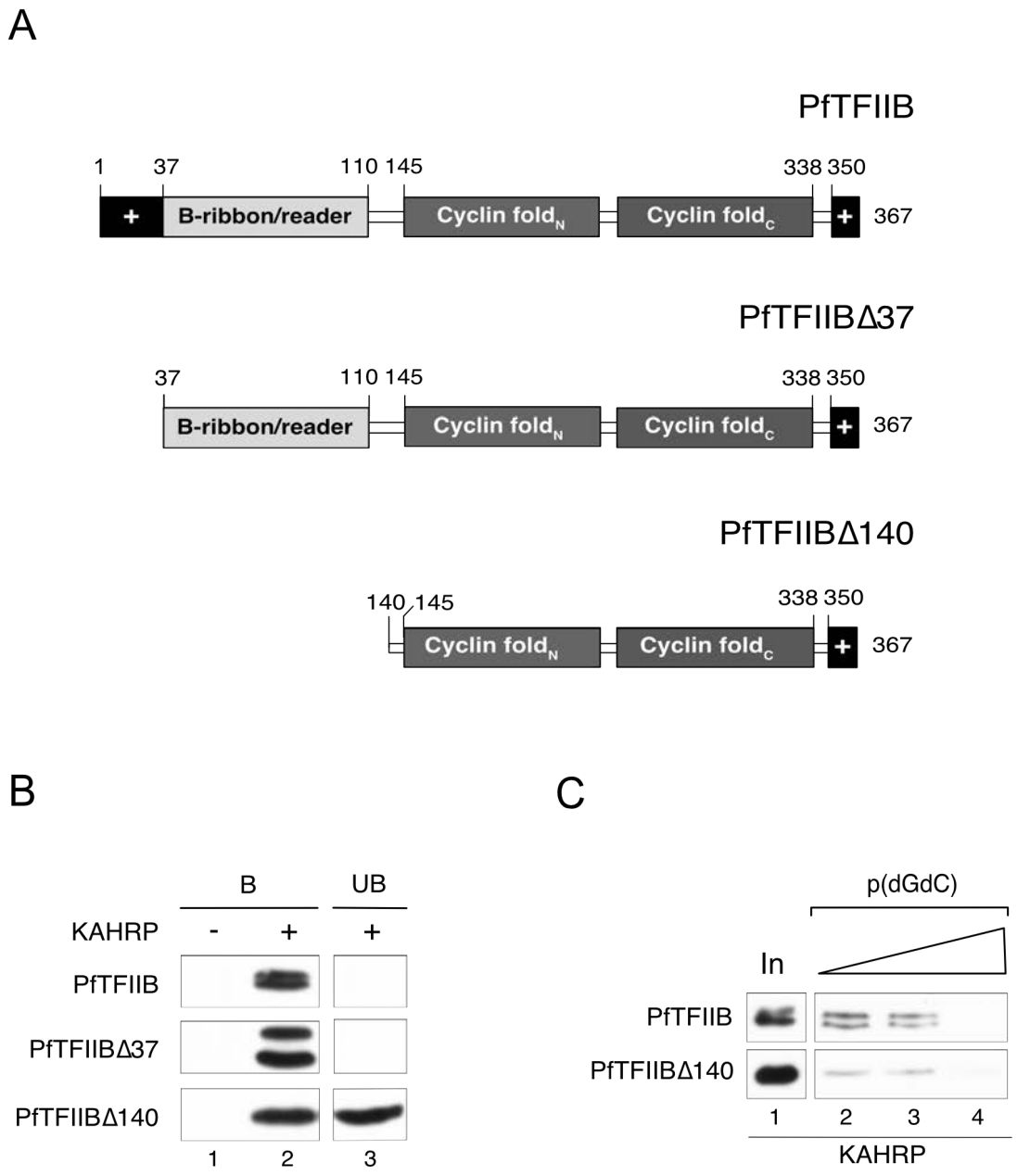

Figure 17. PfTBP-independent DNA-binding activity of full length PfTFIIB and PfTFIIB deletion mutants.

A. Schematic of the PfTFIIB variants used in this study.

B. Immunoblot analysis of PfTFIIB variants binding to immobilised *kahrp* promoter. DNA-binding reactions (40µl) contained 40ng 6His:PfTFIIB, 6His:PfTFIIBΔ37 or 6His:PfTFIIBΔ140 and 1pmol of immobilised *kahrp* promoter DNA (+) or free beads (-), as indicated. Equivalent amounts of unbound (**UB**; lane 3) and bound fractions (**B**; lane 2) were analysed by immunoblot with α-PfTFIIB antibody (1:1000). X-ray film exposure: 30s.

C. Immunoblot analysis of poly(dG-dC) competition of PfTFIIB binding to *kahrp* promoter. DNA-binding reactions (40µl) contained 40ng 6His:PfTFIIB or 6His:PfTFIIBΔ140 and 1pmol of immobilised *kahrp* promoter DNA, 1µg (lane 2), 1.25µg (lane 3), 1.5µg (lane 4) of poly(dG-dC). Bound fractions (1/2th of total elution volume) were analysed by immunoblot with affinity-purified α-PfTFIIB antibody (1:1000). X-ray film exposure: 5min.

experimental probes, a 60bp region of the *kahrp* probe, as well as a 60bp upstream region of the highly variable (*var*) genes were used as an alternative (see Chapter 2.5). On longer probes, it seemed likely that the retention of protein-DNA complexes was caused by multiple binding events, therefore, the smaller probe size was expected to allow the complexed DNA to enter the gel matrix.

As seen from Figure 18, 6His:PfTFIIBΔ140 and 6His:PfTFIIBΔ37 formed stable complexes with the 60bp *var* as well as the *kahrp* probe, in the presence of 100ng poly(dG-dC) competitor

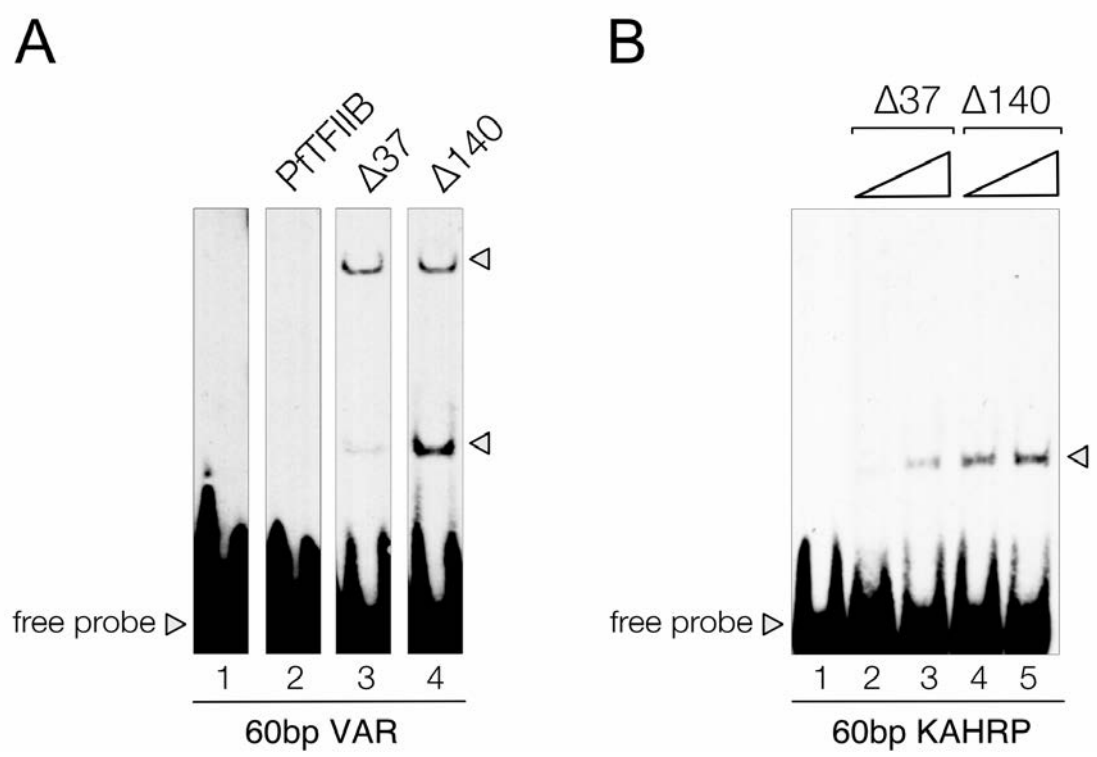

Figure 18. PfTBP-independent binding of PfTFIIBΔ37 and PfTFIIBΔ140 in polyacrylamide EMSA.
A. Analysis of PfTFIIB binding to a 60bp *var* promoter by PAGE EMSA.
10µl binding reactions contained 2fmol 60bp *var* promoter, 100ng poly(dG-dC) and 20ng 6His:PfTFIIB, 20ng 6His:PfTFIIBΔ140 or 10ng 6His:PfTFIIBΔ37. Protein-DNA complexes were resolved in 5% (29:1) native polyacrylamide gel, transferred to a nitrocellulose membrane and visualised by chemiluminescent detection. X-ray film exposure: 30min. Lanes 1-4 originate from the same gel.
B. Analysis of PfTFIIB binding to the *kahrp* promoter by PAGE EMSA.
10µl binding reactions contained 2fmol 60bp *kahrp* promoter, 100ng poly(dG-dC) and 10ng/ 20ng 6His:PfTFIIBΔ37 (lanes 2,3, respectively) or 20/40ng 6His:PfTFIIBΔ140 (lanes 4,5, respectively). Protein-DNA complexes were resolved in 5% (37:1) native polyacrylamide gel, transferred to a nitrocellulose membrane and visualised chemiluminescent detection. X-ray film exposure: 3hr.

DNA. Interestingly, PfTFIIB deletion mutants formed two distinct shifts with the *var* probe, whereas a single complex was seen with the *kahrp* probe. This might indicate the presence of two unique binding sites or formation of two structurally distinct complexes on the *var* probe. At equimolar quantities of 6His:PfTFIIBΔ140 and 6His:PfTFIIBΔ37, significantly higher binding activity with Δ37 was observed, consistent with the findings in ITAs (data now shown). No complexes were observed with full length PfTFIIB (Fig. 18A). Unstable protein-DNA complexes and non-specific DNA binding are known to be sensitive to electrophoresis conditions. The N-terminal basic region of full length PfTFIIB might contribute to non-specific binding of DNA, or is interacting extensively with the competitor DNA, resulting in the inability to form stable complexes with the target DNA. These findings were consistently reproduced in subsequent PAGE EMSAs and provide further evidence for the intrinsic DNA-binding activity of PfTFIIB under specified conditions.

3.5.3. Temperature dependence of PfTFIIB DNA-binding

In order to test the effect of incubation temperature on PfTFIIB-DNA complex formation, 6His:PfTFIIBΔ140 binding reactions were incubated on ice, at room temperature, and at 30˚C in PAGE EMSA. 6His:PfTFIIBΔ140-*kahrp* complex was seen with binding reactions incubated on ice or room temperature, whereas very little complex was detected when incubation was carried out at 30˚C (Fig. 19A, compare lanes 2-4). Interestingly, the total amount of probe detected appeared to be in direct correlation with the incubation temperature. High incubation temperature could lead to partial protein denaturation, and unspecific binding and aggregation could cause the complete inability of the protein-DNA complexes to enter the gel matrix in polyacrylamide EMSAs.

The temperature-dependence of PfTFIIB DNA-binding was further assessed in agarose EMSA conditions. For that purpose, 80ng of 6His:PfTFIIBΔ140 was incubated with *kahrp* promoter at 30˚C and on ice. PfTFIIB-DNA complex was seen migrating directly above the free probe, with approximately 2-fold more probe bound upon incubation on ice compared to 30˚C (Fig. 19B, compare lanes 2 and 5). In order to assess the stability of the complexes, poly(dG-dC) competitor DNA was included in the binding reactions. In the presence of 100ng poly(dG-dC)

competitor DNA, complexes formed at 30˚C were completely disrupted, whereas complex formation was still detectable after incubation on ice (Fig. 19B, compare lanes 3 and 6). Taken together, these results indicate a higher DNA-binding activity of PfTFIIB and higher stability of the resulting complexes at a low temperature, consistent with PfTBP findings (see Chapter 3.2.2). In addition, these results further confirm DNA-binding activity of PfTFIIB, that can be resolved in both agarose and polyacrylamide EMSAs.

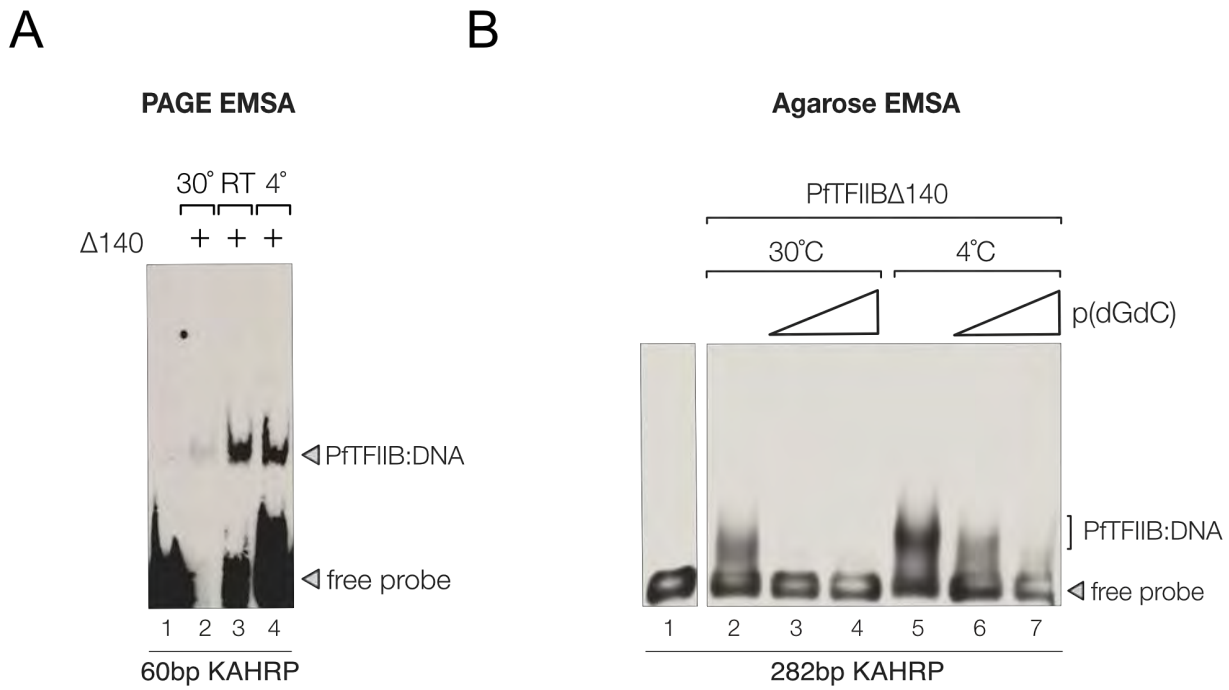

Figure 19. PfTFIIB-DNA complex formation is temperature-dependent.
A. Analysis of PfTFIIBΔ140 binding to the *kahrp* promoter by native PAGE EMSA.
Binding reaction (10µl) contained 2fmol 60bp *kahrp* probe, 120ng 6His:PfTFIIBΔ140 (Δ140) and 50ng poly(dG-dC). Binding reactions were incubated at 30˚C, room temperature (RT) or on ice (4˚C) for 45min, as indicated. Protein-DNA complexes were resolved in 5% (37:1) native polyacrylamide gel, transferred to a nitrocellulose membrane and visualised by chemiluminescent detection. X-ray film exposure: 2hr.
B. Analysis of PfTFIIBΔ140 binding to the *kahrp* promoter by Mg^{2+}agarose EMSA.
Binding reaction (10µl) contained 20fmol 282bp *kahrp* probe, 80ng 6His:PfTFIIBΔ140 and 100ng (lanes 3,6) or 400ng (lanes 4,7) of poly(dG-dC). Binding reactions were incubated at 30 ˚C or on ice (4˚C) for 45min, as indicated. Protein-DNA complexes were resolved in 1.4% agarose gels (0.5xTBE; 5mM $MgCl_2$), vacuum-transferred to a nitrocellulose membrane and visualised by chemiluminescent detection. X-ray film exposure: 2s.

3.5.4. PfTFIIB N-terminus contributes to the PfTBP-independent DNA-binding activity

Next, the binding activities of the full length PfTFIIB and PfTFIIB deletion mutants were compared. The relative protein concentration in each preparation was normalised by SDS PAGE and immunoblot (Fig. 16, data not shown). Equivalent molar amounts of full length PfTFIIB and PfTFIIB deletion mutants were incubated with *kahrp* promoter in the absence of competitor DNA. The probe was quantitatively bound by 11nM of full length 6His:PfTFIIB, whereas no significant binding was observed with 6His:PfTFIIBΔ37 or 6His:PfTFIIBΔ140 at the same molar concentration (Fig. 20 compare lanes 3, 6 and 9). 23nM of 6His:PfTFIIBΔ37 and 71nM of 6His:PfTFIIBΔ140 was minimally required for the detection of DNA-binding (Fig. 20, lanes 7, 11). With the double affinity chromatography purified full length 6His:PfTFIIB:GST, similar complexes to those formed with 6His:PfTFIIB were seen, with the probe quantitatively bound at similar molar concentrations. Since the DNA-binding activity of the full length PfTFIIB is clearly higher than that of the deletion mutants', it seems evident that the 37-residue N-terminal

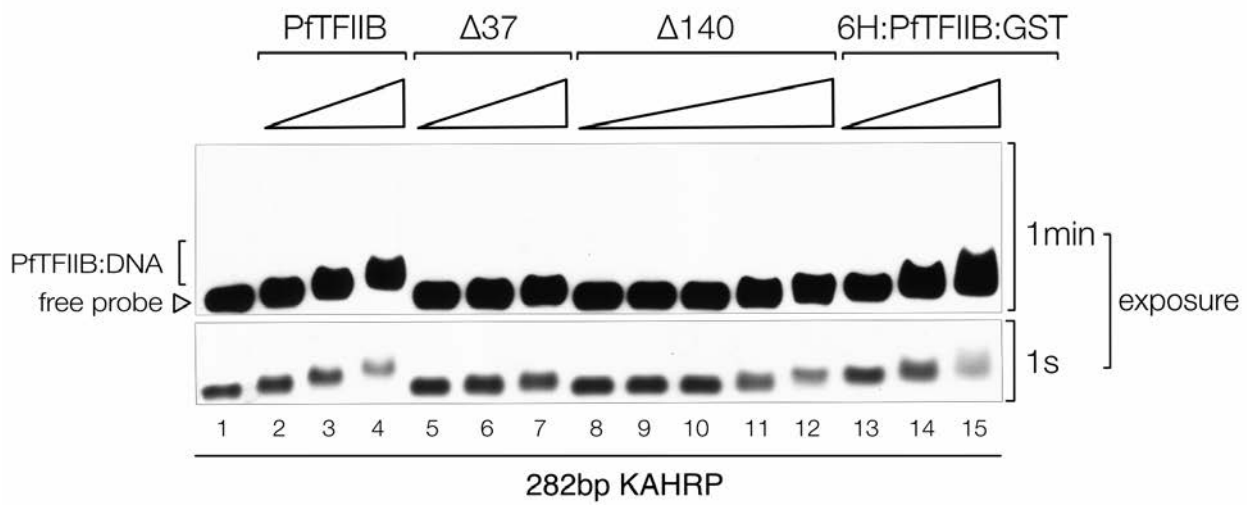

Figure 20. PfTFIIB DNA-binding is modulated by the N-terminal 37-residue region.
Analysis of the binding of full length PfTFIIB and PfTFIIB deletion mutants to the *kahrp* promoter by Mg^{2+} agarose EMSA.10µl binding reactions contained 5fmol 282bp *kahrp* probe, 4.5nM, 11nM, 23nM 6His:PfTFIIB (lanes 2-4, respectively), 6His:PfTFIIBΔ37 (lanes 5-7, respectively), 6His:PfTFIIBΔ140 (lanes 8-10) and 6His:PfTFIIB:GST (lanes 13-15, respectively). 6His:PfTFIIBΔ140 was further titrated up to 36nM and 71nM (lanes 11,12, respectively). Protein-DNA complexes were resolved in 1.4% agarose gels (0.5xTBE; 5mM $MgCl_2$), vacuum-transferred to a nitrocellulose membrane and visualised by chemiluminescent detection. The two panels indicate 1s and 1min exposure of the same experiment.

region contributes to PfTFIIB DNA-binding to a large extent in this assay. The N-terminal region lacking the 37-residues and the C-terminal core domain also have DNA-binding properties under the conditions, albeit to a lesser extent. Taken together, these results are consistent with the binding activity observed in ITAs, and further confirm the intrinsic binding activity of PfTFIIB.

3.5.5. PfTFIIB has sequence preference for A/T-rich sequences

The PfTBP-independent DNA-binding activity of PfTFIIB could imply a specific function of PfTFIIB, such as promoter recognition. Investigation of DNA sequence preference could provide the first evidence for specific DNA-binding of PfTFIIB. Initial assessment of sequence preference was carried out by comparison of PfTFIIB binding to the *P. falciparum kahrp* promoter or the model promoter Ad2ML. The basepair composition differs radically between the two promoters, with guanine-cytosine-(G/C-) content of the *kahrp* promoter sequence being only 23% compared to the 61% of the Ad2ML promoter.

Gradual complex formation with Ad2ML was observed and the probe was quantitatively bound only at the highest amount of PfTFIIB (Fig. 21, lanes 6-8). However, the *kahrp* promoter was

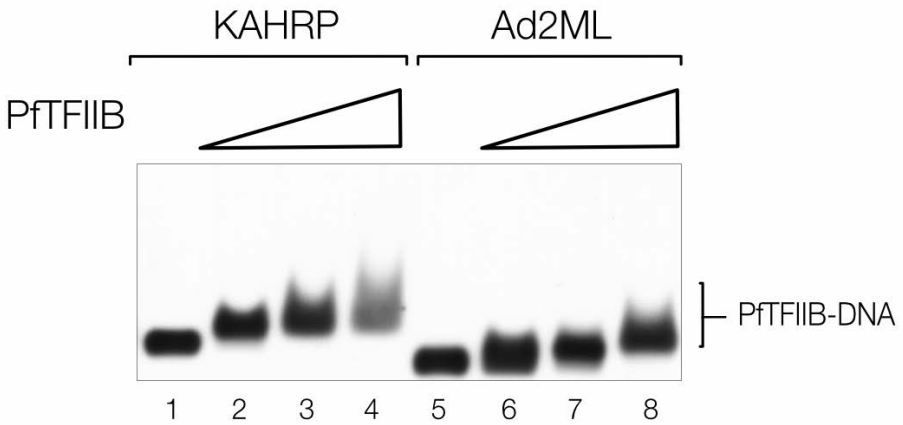

Figure 21. Analysis of PfTFIIB binding to the *kahrp* and Ad2ML promoters by Mg^{2+}agarose EMSA.
Binding reactions (10µl) contained 5fmol 282bp *kahrp* or Ad2ML probe and 10, 20, 40ng of 6His:PfTFIIB:GST (lanes 2-4, 6-8). Protein-DNA complexes were resolved in 1.4% agarose gels (0.5xTBE; 5mM $MgCl_2$), vacuum-transferred to a nitrocellulose membrane and visualised by chemiluminescent detection. X-ray film exposure: 1s.

quantitatively shifted with the lowest amount of PfTFIIB (Fig. 21, lane 2). Further mobility shift of the probe upon increasing the PfTFIIB concentration was observed with the *kahrp* promoter only, indicating binding at multiple locations with low affinity. These differences suggest that PfTFIIB binds A/T-rich sequences with higher affinity.

To further investigate if PfTFIIB has a preference for a consensus TATA-motif, PfTFIIB-*kahrp* complexes were challenged with up to 20-fold molar excess of unlabelled 126bp TATA-box containing, or TATA-less mouse TdT (mTdT) promoter. Interestingly, with 2-fold molar excess of the TATA-containing competitor, the PfTFIIB-DNA complex was disrupted and an alternative complex appeared, which migrated between the free probe and the original PfTFIIB-*kahrp* complex (Fig. 22, compare lanes 2 and 15). Increasing amounts of the TATA-containing competitor did not further affect the complex formation (Fig. 22, compare lanes 16, 17). A similar intermediate complex was seen with the TATA-less competitor only at 20-fold excess of the competitor over the *kahrp* probe (Fig. 22, compare lanes 2, 11 and 14). Free probe was not fully recovered with either competitor at these concentrations. It seems, PfTFIIB binds the 282bp *kahrp* promoter at multiple locations with different affinity, and the binding of PfTFIIB to some of those locations is more sensitive to the TATA-containing sequence than the TATA-less sequence. Further competition experiments are needed to confirm this finding.

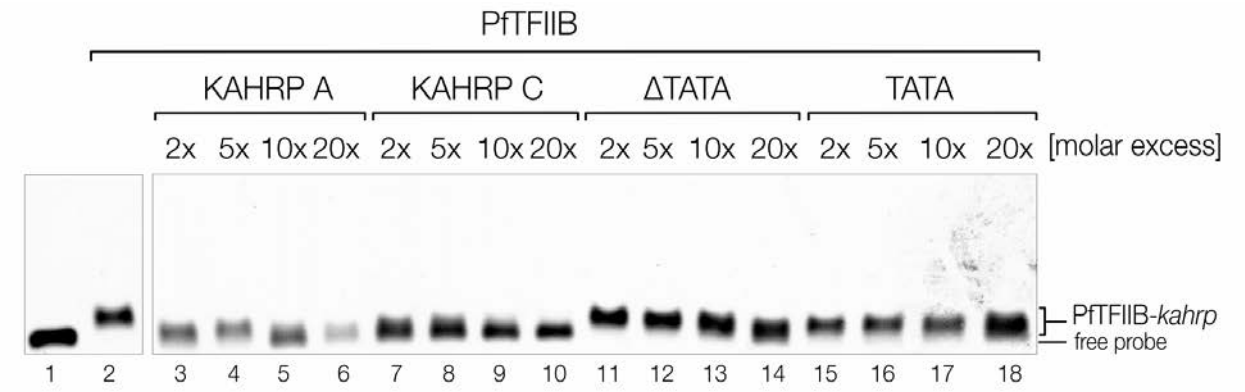

Figure 22. PfTFIIB binds preferentially A/T-rich sequences.
PfTFIIB DNA-binding template competition analysis by Mg^{2+} agarose EMSA.
Binding reactions (10µl) contained 5fmol 282bp *kahrp* probe, 20ng 6His:PfTFIIB:GST, 2,5,10 or 20x molar excess of unlabelled 138bp *kahrp* 5' region (*kahrp* A), 157bp *kahrp* 3' region (*kahrp* C), TATA-less (ΔTATA) or TATA-containing (TATA) mTdT promoters. Protein-DNA complexes were resolved in 1.4% agarose gels (0.5xTBE; 5mM $MgCl_2$), vacuum-transferred to a nitrocellulose membrane and visualised by chemiluminescent detection. X-ray film exposure: 10s.

To further identify if a specific region of the *kahrp* promoter is preferentially bound by PfTFIIB, PfTFIIB-*kahrp* complexes were challenged with up to 20-fold molar excess of unlabelled *kahrp* 5' region spanning nucleotides 1-138 (*kahrp* A) as well as a 157 bp 3' region spanning nucleotides 125-282 (*kahrp* C). This approach allows an assessment of the binding of PfTFIIB to each half of the 282bp *kahrp* probe. As seen in Figure 22, free probe was fully recovered with 20-fold excess of *kahrp* C (Fig. 22, lane 10) and *kahrp* A (Fig. 22, lane 6). Further experiments confirmed complete disruption with either *kahrp* competitor at minimally 20-fold molar excess (data not shown).

As free probe was fully recovered only when the complex was challenged with 20X *kahrp* promoter and intermediate complex persisted with the TATA variants, the *kahrp* promoter either harbours a specific binding motif distinct from the consensus in the mTdT probe, or the preferred binding of PfTFIIB to *kahrp* promoter is the result of its overall higher A/T-content.

3.6. Cooperative binding of PfTBP and PfTFIIB

3.6.1. PfTBP recruits PfTFIIB on *kahrp* and *gbp-130* promoter regions

In eukaryotes, TFIIB interacts with promoter-bound TBP and stabilises the TBP-promoter complex. In order to first assess interactions between PfTFIIB and PfTBP, immobilised template assays were performed, where PfTBP was incubated in the presence and absence of PfTFIIB and immobilised *kahrp* and *gbp-130* promoters. It was previously observed, that PfTFIIB elicits extensive DNA-binding activity in ITAs (see Chapter 3.5). In order to detect an effect on PfTFIIB promoter-binding in the presence of PfTBP, PfTFIIB DNA-binding was suppressed with 1.5µg of poly(dG-dC). As seen in Figure 23A, little to no residual PfTFIIB binding to the *kahrp* promoter was detected at these concentrations of the competitor, consistent with previous results (see Fig. 17).

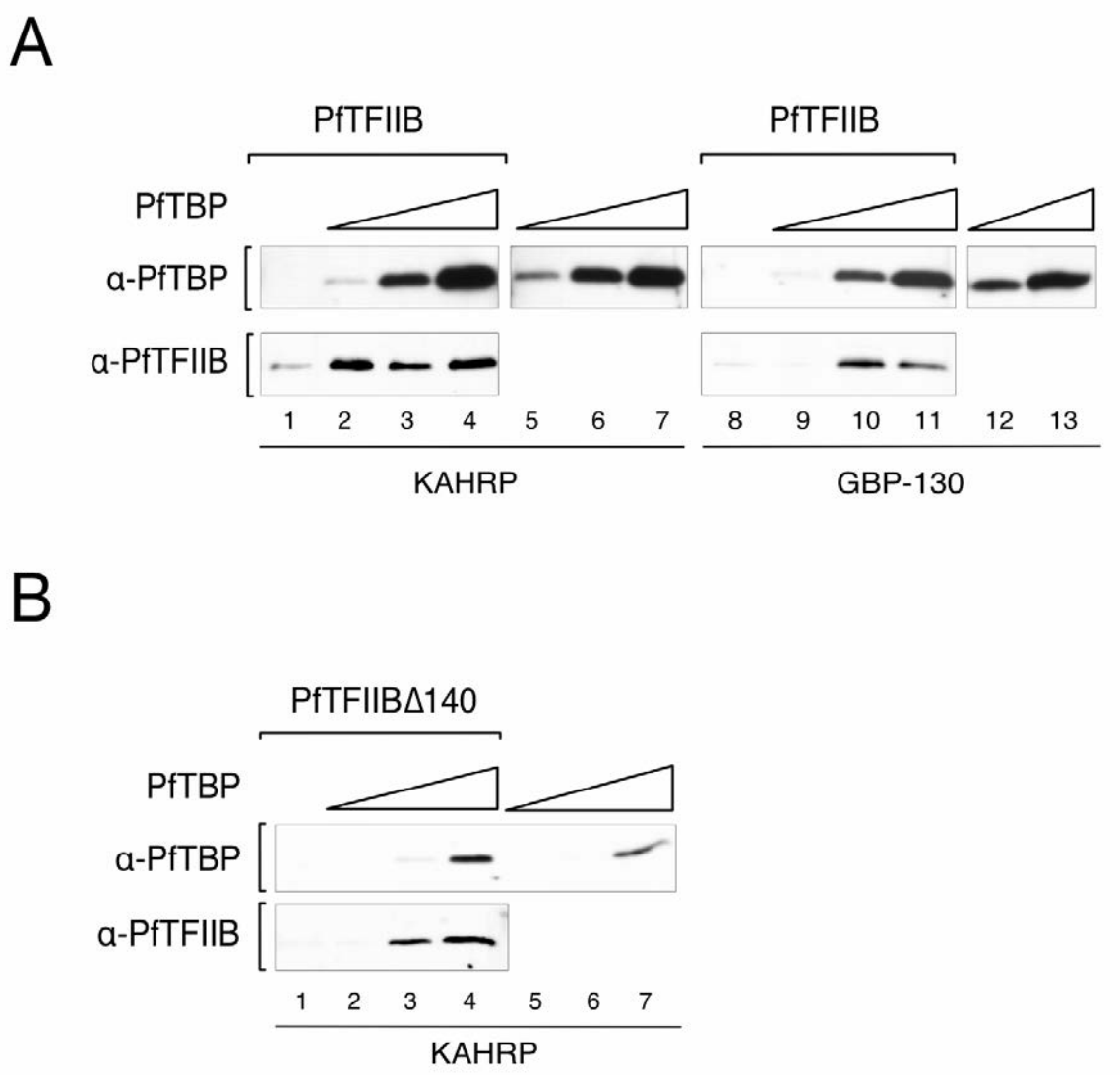

Figure 23. PfTBP recruits PfTFIIB to putative *P. falciparum* promoter regions.
A. Immunoblot analysis of PfTFIIB binding to immobilised *gbp-130* and *kahrp* promoters in the presence of PfTBP. DNA-binding reactions (40µl) contained 40ng 6His:PfTFIIB and 12ng (lanes 2,9), 30ng (lanes 3,10), 60ng (lanes 4,11) 6His:PfTBPc, as indicated, 1pmol of immobilised *kahrp* (lanes 1-7) or *gbp-130* (lanes 8-13) promoter DNA and 1.5µg of poly(dG-dC). Bound fractions ($1/2^{th}$ total volume) were analysed by immunoblot with α-PfTBP rabbit antisera (1:500) or α-PfTFIIB antibody (1:1000), as indicated. X-ray film exposure: 10min.
B. Immunoblot analysis of PfTFIIBΔ140 binding to immobilised *kahrp* promoter in the presence of PfTBP.
DNA-binding reactions (40µl) contained 40ng 6His:PfTFIIBΔ140 and 12ng (lanes 2,5), 30ng (lanes 3,6), 60ng (lanes 4,7) 6His:PfTBPc, as indicated, 1pmol of immobilised *kahrp* promoter DNA and 1.5µg of poly(dG-dC). Bound fractions ($1/2^{th}$ total volume) were analysed by immunoblot with α-PfTBP rabbit antisera (1:500) or α-PfTFIIB antibody (1:1000), as indicated. X-ray film exposure: 10min.

As shown in Figure 23A, PfTFIIB binding to *kahrp* and *gbp-130* promoters was significantly stimulated by the presence of PfTBP (lower panel, compare lanes 1 and 2, 8 and 10). Increasing amounts of PfTBP did not appear to further stimulate PfTFIIB binding. As previously noted (Chapter 3.5.1, Fig. 17), the binding activity of PfTFIIBΔ140 is at least 2-fold weaker compared to full length PfTFIIB, and residual binding to *kahrp* was nearly undetectable (Fig. 23B, lower panel, lane 1). Similarly to full length PfTFIIB, PfTFIIBΔ140 binding to the promoter was greatly stimulated by the presence of PfTBP (Fig. 23B, lower panel, compare lanes 1 and 3). In contrast, PfTBP DNA-binding appeared unaffected by the presence of full length PfTFIIB or PfTFIIBΔ140 (Fig. 23A compare lanes 5-6 and 2-4, panel B compare lanes 4-7 and 2-4). Taken together, these results provide strong evidence for the recruitment of PfTFIIB to the *kahrp* and *gbp-130* promoters by PfTBP, even in the presence of high excess of poly(dG-dC) competitor. Furthermore, in agreement with previous studies in other eukaryotic systems (Nikolov et al., 1995), the C-terminal domain of PfTFIIB is sufficient for the recruitment by PfTBP. These results clearly indicate specific interactions between the transcription factors, and the potential for PfTBP-PfTFIIB-promoter complexes were investigated further in EMSAs.

3.6.2. Formation of ternary PfTBP-PfTFIIB-promoter complex

In eukaryotic model systems, TFIIB stabilisation of the TBP-promoter complex is evident from the enhancement of TBP binding to the promoter in the presence of TFIIB, resulting in TBP-TFIIB-promoter ternary complexes in EMSAs. In order to assess the potential for PfTBP-PfTFIIB-promoter complex formation in EMSAs, the independent binding of PfTFIIB or PfTBP was suppressed by 400ng of poly(dG-dC) competitor DNA. Under these conditions, no significant binding of PfTBP to *kahrp* promoter was seen. Similarly to PfTBP, no significant binding of PfTFIIB was observed, with some smearing indicative of residual binding (Fig. 24). PfTBP incubation in the presence of PfTFIIB resulted in clear decrease in the probe mobility, however the complex could not be fully resolved under these conditions. The majority of independent promoter-binding of PfTFIIBΔ37 was also abolished under the conditions, as seen in Figure 24 (lane 6). However, a distinct complex was observed with PfTBP in the presence of PfTFIIBΔ37. The clear change in mobility and increased probe binding further indicates that the

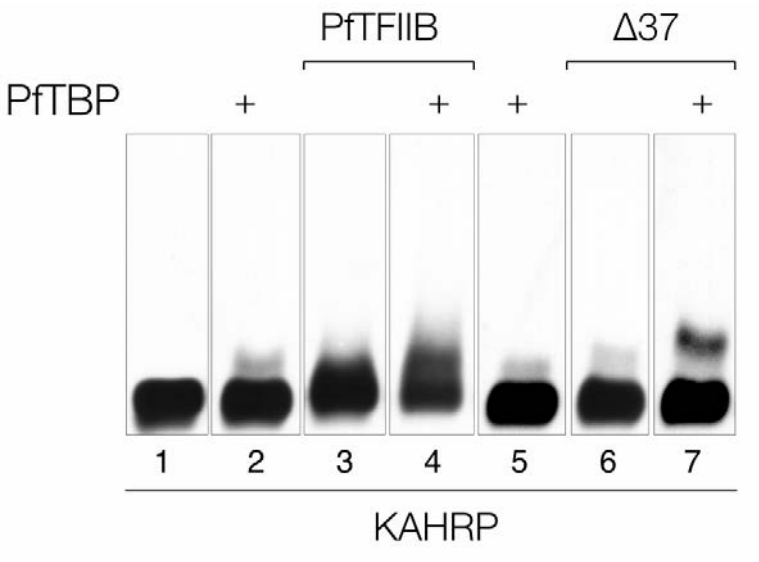

Figure 24. Distinct complex formation with PfTBP in the presence of PfTFIIB.
Analysis of PfTBP binding to the *kahrp* promoter in the presence of PfTFIIB by Mg^{2+} agarose EMSA.
Binding reactions (10µl) contained 5fmol 282bp *kahrp* probe, 3ng 6His:PfTBP (lanes 2, 4-5, 7), 40ng 6His:PfTFIIB:GST (lanes 3-4) or 60ng 6His:PfTFIIBΔ37 (lanes 6-7), as indicated, and 400ng poly(dG-dC). Protein-DNA complexes were resolved in 1.4% agarose gels (0.5xTBE; 5mM $MgCl_2$), vacuum-transferred to a nitrocellulose membrane and visualised by chemiluminescent detection. X-ray film exposure: 10s.

complex observed is a new unique complex, and not the combined effect of individual interactions of PfTFIIB and PfTBP with the promoter. PfTFIIBΔ140 in conjunction with PfTBP did not induce significant complex formation (data not shown). These results indicate the formation of a distinct complex in the presence of both PfTBP and PfTFIIB and are in agreement with the known model of TFIIB enhancement of TBP-DNA complex formation.

Next, the presence of both PfTFIIB and PfTBP in the newly formed complex was determined. For that purpose, purified α-PfTBP or α-PfTFIIB antibodies were included in the binding reactions. As seen from Figure 25A, addition of α-PfTBP antibody resulted in the complete disruption of the PfTBP-PfTFIIBΔ37-*kahrp* complex. This suggests that the binding of the antibody to PfTBP obstructs the PfTBP interactions with DNA and/or PfTFIIB and thus confirms the presence of PfTBP in the complex. In contrast, addition of α-PfTFIIA antibody which was produced and purified under identical conditions did not affect the complex. Similarly to α-PfTBP, α-PfTFIIB disrupted the protein-DNA complex, indicating the presence of PfTFIIB (Fig. 25B, compare lanes 4, 6). α-PfTBP and α-PfTFIIB did not affect the free probe

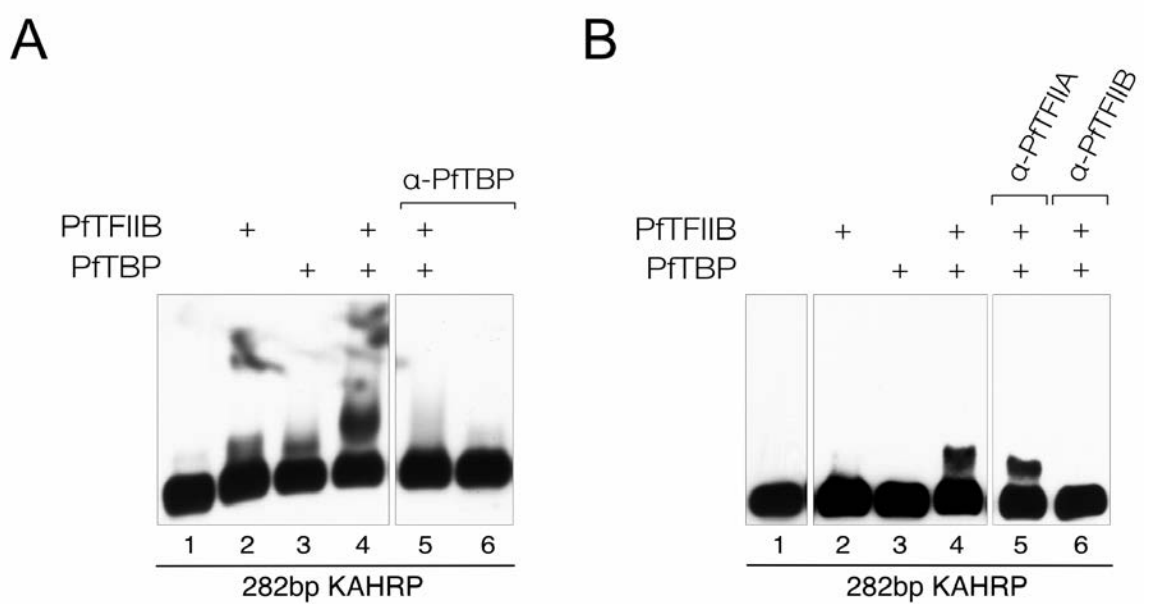

Figure 25. Assembly of a PfTBP-PfTFIIB-DNA ternary complex on *P. falciparum kahrp* promoter region.
Analysis of PfTBP-PfTFIIB-*kahrp* complex formation by Mg^{2+} agarose EMSA.
A. Binding reactions (10µl) contained 5fmol 282bp *kahrp* probe, 400ng poly(dG-dC), 12ng 6His:PfTBPc, 60ng 6His:PfTFIIBΔ37 and 1µg α-PfTBP, as indicated. Protein-DNA complexes were resolved in 1.4% agarose gels (0.5xTBE; 5mM $MgCl_2$), vacuum-transferred to a nitrocellulose membrane and visualised by chemiluminescent detection. X-ray film exposure: 10s.
B. Binding reactions (10µl) contained 5fmol 282bp *kahrp* probe, 400ng poly(dG-dC), 3ng 6His:PfTBPc, 60ng 6His:PfTFIIBΔ37 and 1µg α-PfTFIIA or 200ng α-PfTFIIB, as indicated. Protein-DNA complexes were resolved in 1.4% agarose gels (0.5xTBE; 5mM $MgCl_2$), vacuum-transferred to a nitrocellulose membrane and visualised by chemiluminescent detection. X-ray film exposure: 1min.

(Fig. 25A, lane 6; data not shown). Together, these results confirm that both PfTFIIB and PfTBP are present in the complex. Taken together, these findings provide strong evidence for the formation of a *bona fide* PfTBP-PfTFIIB-DNA complex on putative *P. falciparum* promoter region.

3.6.3. Sequence preference of PfTBP-PfTFIIB-promoter complex

Next, the sequence specificity of the PfTBP-PfTFIIB-*kahrp* ternary complex was analysed. For this purpose, pre-assembled PfTBP-PfTFIIB-*kahrp* complexes were challenged with 2-20x

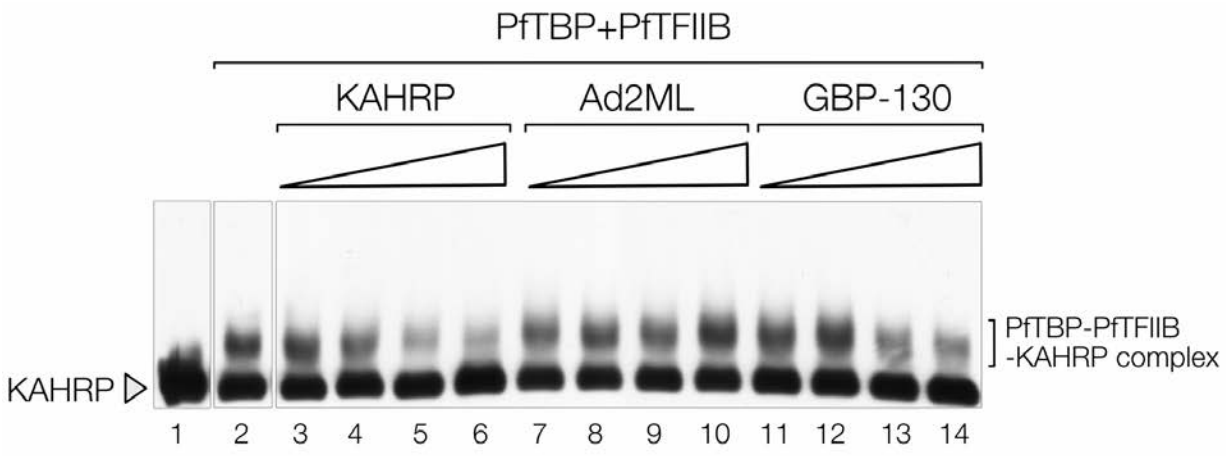

Figure 26. Sequence specificity of the PfTBP-PfTFIIB-*kahrp* complex.
PfTBP-PfTFIIB-*kahrp* template competition analysis by Mg^{2+} agarose EMSA.
Binding reactions (10µl) contained 5fmol 282bp *kahrp* probe, 12ng 6His:PfTBPc, 60ng 6His:PfTFIIBΔ37, 2x (lanes 3, 7, 11), 5x (lanes 4, 8, 12),10x (lanes 5, 9, 13) or 20x (lanes 6, 10, 14) molar excess of unlabelled 282bp *kahrp, gbp-130* or 297bp Ad2ML promoters, as indicated. Protein-DNA complexes were resolved in 1.4% agarose gels (0.5xTBE; 5mM $MgCl_2$), vacuum-transferred to a nitrocellulose membrane and visualised by chemiluminescent detection. X-ray film exposure: 5s.

---

molar excess of unlabelled 282bp *kahrp*, *gbp-130* or 297bp Ad2ML promoters. As seen in Figure 26, the PfTBP-PfTFIIB-*kahrp* complex decreased in the presence of excess unlabelled *kahrp* or *gbp-130* promoters, but remained largely unaffected in the presence of excess Ad2ML promoter. Previous findings suggest similar binding affinity of PfTBP to *kahrp* and Ad2ML promoters (Fig. 10), whereas a clear preference towards *kahrp* promoter was seen with PfTFIIB (Fig. 21). Since the PfTBP-PfTFIIB-*kahrp* complex appeared to be inert towards Ad2ML promoter, it is apparent that the sequence specificity of the PfTBP-PfTFIIB-DNA complex in comparison to individual PfTBP binding had changed.

4. Discussion

4.1. Expression and purification of PfTBP

The expression of recombinant *Plasmodium* proteins is known to be challenging, often resulting in the production of truncated proteins or low yield (reviewed in Birkholtz et al., 2008). Adverse effects on the host cells upon recombinant expression of plasmodial proteins have been reported previously (Vedadi et al., 2007). Consistent with these reports, apparent toxicity of *P. falciparum* proteins upon over-expression in *E. coli* has consistently been observed within our research group, resulting in inhibition of host cell growth. Therefore, recombinant PfTBPc in this study was generated from uninduced cultures. Purification of PfTBPc was a multi-step approach using Ni-affinity chromatography followed by ion exchange chromatography. Early PfTBPc elution fractions from Ni-NTA agarose were found to have lower DNA-binding activity than late fractions, potentially due to aggregation of PfTBPc in the early fractions that contained PfTBPc at higher concentrations. Higher salt concentrations during Ni-NTA agarose purification appeared to alleviate aggregation of PfTBPc, therefore cleared lysate preparation and Ni-NTA agarose purification at higher salt concentration (500mM KCl instead of 300mM KCl) could streamline the process and eliminate the need for re-purification by Ni-NTA beads. The Ni-affinity purified PfTBPc bound to strong cation exchange resin as expected at its pI of 9.3, and eluted at high salt concentrations. The purity of the resulting PfTBPc preparation was deemed sufficient for subsequent characterisation and the protocol established is therefore proposed to serve as a standard for PfTBP purification.

4.2. Initial characterisation of PfTBP in DNA binding assays

The ultimate goal of this research project is to investigate the DNA-binding characteristics of the key protein factors involved in transcription initiation in *P. falciparum*, in order to identify *P. falciparum* core promoter regions and the DNA elements directing transcription initiation complex assembly. The DNA-binding activities of PfTBP and PfTFIIB were investigated using two independent approaches - electrophoretic mobility shift assays (EMSAs) and immobilised template assays (ITAs). Due to the scarcity of current knowledge about optimal transcription

conditions for the parasite, initial assays were set up utilising the conditions of the well-established human model system (Ge et al., 1996, Malecova et al., 2007).

From the DNA binding assays, three distinct observations were made regarding the DNA-binding characteristics of isolated PfTBPc. Firstly, PfTBP forms a fast-migrating complex with the tested promoter regions in Mg^{2+} agarose EMSAs and, upon increasing concentrations, the mobility of the complex decreases further. Secondly, PfTBP appears to bind to Ad2ML and putative *P. falciparum* promoters with similar affinity, as observed in ITA and EMSA. Thirdly, the stability of the PfTBP-DNA complex is significantly enhanced by lowering the incubation temperature.

In eukaryotes, TBP binding to the TATA box is a two-stage process, where initial TBP binding to the TATA box does not affect the target DNA, followed by a slow transition into a fully bent complex that is subject to stabilisation by TFIIB and TFIIA (Hoopes et al., 1992, Zhao and Herr, 2002). The DNA-binding activity of TBP resides in the highly conserved C-terminal domain, that is a saddle-shaped molecule consisting of two direct repeats of five β-sheets and two α-helices (Nikolov et al., 1996). TBP interacts with the DNA minor groove with its concave surface, where most notably two phenylalanine (Phe) residues insert into the 5’ and 3’ end of the TATA box (F288 and F197 in humans, respectively), inducing a kink in the DNA(Nikolov et al., 1996). Two additional phenylalanine residues (F305 and F214 in humans) support the inserted Phe residues and stabilise the induced kinks (Nikolov et al., 1996). Collectively, these interactions induce a severe bend in the DNA of about ~80˚, whereby the bending angle is dependent on the depth at which the Phe residues are inserted between the base pairs in the DNA (Starr et al., 1995, Nikolov et al., 1996). In electrophoretic mobility shift assays, the mobility of protein-DNA complex relative to free DNA is determined by the molecular weight and tertiary (or quaternary) structure of the DNA-bound protein(s), as well as the conformation of the DNA. DNA bending significantly reduces the mobility of protein-DNA complexes in EMSA gels. Therefore, due to the extreme TBP-induced bending of DNA, TBP-DNA complexes are unusually slow-migrating (Horikoshi et al., 1990, Maldonado et al., 1990). In contrast, protein-DNA complexes lacking major distortion in DNA structure are highly mobile in Mg^{2+} agarose EMSAs, and are seen migrating in close proximity to the free probe even when multiple proteins

are bound (Lieberman and Berk, 1994, Malecova et al., 2015).

Interestingly, *P. falciparum* TBP amino acid sequence analysis shows that one of the two Phe residues, that insert between the DNA and are highly conserved between eukaryotic TBPs, is replaced by isoleucine (Ile183) (Thomas Oelgeschläger, unpublished observations). Mutations of single or both Phe residues of the two Phe pairs have been shown to render TBP incapable of forming stable TBP-DNA complexes in eukaryotic and archaeal systems (Klejman et al., 2005, Gietl et al., 2014). The results in this study reveal a fast-migrating PfTBP-DNA complex with *P. falciparum* and Ad2ML promoters, observable directly above the free probe. Considering the absence of one kink-inducing Phe residue in *P. falciparum* TBP and the appearance of a fast-migrating PfTBP-DNA complex, it seems likely, that PfTBP does not induce the degree of DNA-bending observed with other eukaryotes.

TBP has a relatively low DNA-binding specificity compared to typical DNA-binding transcription regulators (Coleman and Pugh, 1995, Weideman et al., 1997) and can form complexes with variable TATA sequences that bear little resemblance to the optimal TATA-motif (5-TATATAAG-3) (Buratowski et al., 1989, Wong and Bateman, 1994). We observed a PfTBP concentration-dependent decrease in the mobility of the DNA probe that can be explained by successive binding of multiple PfTBP molecules to lower affinity binding sites at higher protein concentrations. Thus, PfTBP seems to bind at multiple locations within *P. falciparum* putative *kahrp* promoter regions, as well as the adenovirus 2 major late model promoter.

Seminal studies of human TFIID binding to DNA reported an increased binding of TBP to TATA sequences at 30˚C in comparison to 4˚C (Lee et al., 1992). These conditions are thought to be thermodynamically favourable for DNA bending during TBP-DNA complex formation (Petri et al., 1995). An incubation temperature of 30˚C has since become the standard temperature in DNA binding and transcription assays of human, *Drosophila* and yeast systems (Thomas and Chiang, 2006). Despite the preferred incubation temperature of 30˚C of wild-type yeast TBP (yTBP), yTBP C-terminal domain forms stable complexes with DNA when incubated at 0˚C or 4˚C (Lieberman et al., 1991, Lee et al., 1992, Kuddus and Schmidt, 1993). Similarly to yeast, PfTBP core domain formed stable complexes at lower temperatures. However, in contrast to yTBP, PfTBP DNA-binding was found to be destabilised by higher incubation temperatures.

This feature of PfTBPc was not investigated further in this MSc project. The potential need for optimising binding conditions was anticipated and it is possible that further adjustments of reaction conditions could significantly affect the outcome of future DNA-binding experiments with *P. falciparum* proteins.

4.3. Discrepancies with findings reported by Ruvalcaba-Salazar et al. (2005).

The initial characterisation of *P. falciparum* TBP in this study has revealed several features of PfTBP that were consistently observed throughout the study and evident from highly reproducible experiments. Firstly, recombinant PfTBP expression is achieved from highly specialised cells adapted for the bacterial expression of proteins from AT-rich genomes and IPTG-induction was found to inhibit cell growth. Thus, in agreement with previous studies on recombinant plasmodial protein expression, PfTBP was found to be extremely toxic to the host cells, resulting in low yield of soluble PfTBP in induced cultures. Secondly, despite comprehensive optimisation of native PAGE EMSA conditions, PfTBP-DNA complexes could not be resolved in this assay. Thirdly, we show extensive DNA-binding activity of PfTBP in ITA and Mg^{2+} agarose EMSAs and provide evidence for PfTBP DNA-binding at multiple locations across the tested DNA probes. Lastly, our results suggest similar binding affinity of PfTBP on putative *P. falciparum* promoter regions as well as Ad2ML promoter.

The study conducted by Ruvalcaba-Salazar et al. (2005) reports high expression levels of recombinant PfTBP core domain from IPTG-induced cultures, using *E.coli* DH5α cells. Ruvalcaba-Salazar et al. (2005) demonstrate specific PfTBP-DNA complex formation in native PAGE EMSAs, even at very high (up to 250ng) concentrations of PfTBP. Furthermore, a distinct TATA-box at position -81 on the putative *kahrp* promoter and a TATA-like sequence at position -186 on the *gbp-130* promoter as exclusive PfTBP binding sites on the promoters were identified using native PAGE EMSAs and DNAse I footprinting assays. These findings were validated *in vivo* by ChIP assays, utilising polyclonal antibody against human TBP, that was shown to identify PfTBP from nuclear *Plasmodium* extracts.

Several of these results reported by Ruvalcaba-Salazar et al. (2005) are incompatible with our findings. The high levels of PfTBP expression achieved from IPTG-induced cultures clearly contradict our findings, where IPTG-induced PfTBP expression inhibited cell growth and caused premature cell death. It is therefore surprising, that the recombinant expression of PfTBP was achieved in *E.coli* DH5α cells which are commonly used for routine cloning applications and are not adapted for the expression of *P. falciparum* genes with the extreme codon bias. Remarkably, amid the abundance of potential TATA-motifs in the *gbp-130* and *kahrp* promoter regions, Ruvalcaba-Salazar et al. (2005) identified specific TATA-motifs using native PAGE EMSAs. Our findings in Mg^{2+} agarose EMSAs indicate PfTBP binding at multiple locations on these *Plasmodium* promoter regions over a narrow concentration range. This contrasts strongly the highly selective binding of PfTBP in DNAse I footprinting assays reported by the authors.

In conclusion, there are fundamental differences between the key findings of this MSc research project and the results reported by Ruvalcaba-Salazar's group. These apparent discrepancies require further investigation.

4.4. Expression and purification of 6His:PfTFIIB:GST

Similarly to PfTBPc expression, IPTG-induction of the recombinant expression of 6His:PfTFIIB:GST resulted in reduced cell growth compared to uninduced cultures. However, 6His:PfTFIIB:GST was sufficiently expressed in uninduced cultures for purification and further analysis. This is attributed to leaky expression from the IPTG-inducible *lac*UV5 promoter, that drives T7 polymerase expression. Optimal conditions for the uninduced expression of *P. falciparum* proteins were investigated previously by our group and addition of glucose in the growth media was found to produce higher cell-density and cell mass, and repress initial leaky expression from the *lac*UV5 promoter. The expression from *lac*UV5 is regulated by LacI repressor and catabolite repression. LacI repressor is active in the absence of lactose and can be inactivated by IPTG to induce expression from the *lac*UV5 promoter. In addition, the expression of the *lac* operon is mediated by catabolite activator protein (CAP) which is naturally expressed in *E. coli* and is inactive in the presence of glucose. Glucose depletion from the media activates CAP, allowing for low levels of T7 polymerase expression from the *lac*UV5 promoter and subsequently the expression of PfTFIIB. In conclusion, PfTFIIB expression was achieved by auto-induction was set as a standard for further PfTFIIB expression in this study.

6His:PfTFIIB:GST was purified using double-affinity chromatography. Interestingly, 6His:PfTFIIB:GST purified as a doublet band in SDS PAGE, with apparent molecular weights of 68 and 70kDa. A similar doublet with the same relative migration of the two bands was later seen with 6His-tagged PfTFIIB and PfTFIIBΔ37. Both bands of the doublets observed with these proteins reacted equally with α-PfTFIIB antibody. However, the C-terminal core domain migrated as a single band in SDS PAGE. The PfTFIIB residues 38-82 show high homology to other eukaryotic TFIIBs and are predicted to form a zinc-finger (Steven Bing, MSc thesis 2015). Zinc-finger-containing proteins migrating as a doublet on SDS PAGE have been reported previously (Lindstrom et al., 2014). It is seems plausible that PfTFIIB purifies as a heterogenous population with distinct zinc-finger conformations, that are not fully denatured by SDS. Future studies, utilising point mutations in structurally crucial residues of the predicted zinc-finger could further substantiate these findings.

4.5. Expression and purification of 6His-tagged PfTFIIB and PfTFIIB deletion mutants

The expression conditions established by pilot studies with 6His:PfTFIIB:GST were subsequently utilised for the large scale expression of 6His:PfTFIIB, 6His:PfTFIIBΔ37 and 6His:PfTFIIBΔ140. By utilising the 6His-tag as well as the high pI of all the PfTFIIB variants, a standard purification scheme utilising Ni-NTA affinity and ion exchange chromatography was established.

4.6. PfTBP-independent PfTFIIB DNA-binding activity

Initial DNA-binding experiments indicated strong PfTBP-independent DNA-binding activity of PfTFIIB. In order to determine the PfTFIIB regions with DNA-binding activity, two N-terminal deletion mutants were generated, lacking the highly basic N-terminal residues 1-37, or the whole N-terminus (residues 1-140) of *P. falciparum* TFIIB. Subsequent DNA-binding assays revealed a number of interesting features of PfTFIIB: (i) C-terminal core domain is sufficient to bind DNA in a PFTBP-independent manner and this binding activity is significantly enhanced by the N-

terminus (ii) Similarly to PfTBP, PfTFIIB DNA-binding activity is increased at a lower temperature and resulting complexes are significantly more stable. (iii) Increasing PfTFIIB concentrations corresponded to the resulting decrease in the mobility of the PfTFIIB-DNA complexes in Mg^{2+} agarose EMSAs, indicative of PfTFIIB binding at multiple locations on the probe. (iv) PfTFIIB DNA-binding affinity is higher for *P. falciparum* promoter regions or generally A/T-rich sequences, in comparison to G/C-rich promoter regions containing a consensus TATA-motif.

In eukaryotes, TFIIB binds to promoter-bound TBP and stabilises the resulting TBP-TFIIB-DNA complex via interactions with IIB recognition sequences up- and downstream of the TATA-motif (BREu and BREd, respectively; discussed in detail in Chapter 1.2.4). The TFIIB C-terminal region interacts with TBP and BREs and has been shown to be sufficient for the formation of TBP-TFIIB-DNA ternary complex in crystal structures (Nikolov et al., 1995). The N-terminal region of PfTFIIB is involved in the recruitment of RNAPII and TFIIF to the pre-initiation complex and DNA-binding activity of this region has not been previously reported. *P. falciparum* TFIIB N-and C-termini contain 37- and 17-residue regions, respectively, that are positively charged (pI 11.7 and 10.8, respectively) and absent in human TFIIB. In this MSc research, PfTBP-independent DNA-binding activity was first observed in ITAs, where the PfTFIIB C-terminal core domain was found to be sufficient to bind to the *P. falciparum* putative *kahrp* promoter region. The binding of C-terminal domain was further enhanced by the presence of minimally residues 38-140 of the N-terminus. Consistent with the ITA findings, the more sensitive agarose EMSA experiments indicated that the N-terminal 37-residue basic region of PfTFIIB significantly contributes to IIB binding activity. The C-terminal region alone had lower DNA-binding activity compared to full length PfTFIIB in Mg^{2+} agarose EMSAs. In native PAGE EMSAs, full length PfTFIIB-DNA complexes could not be detected. High binding activity of the N-terminus, potentially due to non-specific binding by the highly charged region, could lead to aggregation of the protein-DNA complexes, which cannot be resolved under the conditions. Taken together, the results presented in this study suggest that a proportion of DNA-binding activity of PfTFIIB can be attributed to the charged N- and C-termini, that bind DNA in a non-specific manner due to charge attraction. However, considering limited structural information on PfTFIIB, specific binding via DNA-binding motifs cannot be excluded.

In Mg^{2+} agarose EMSAs, PfTFIIB formed fast-migrating complexes with DNA that are different from the previously observed PfTBP-DNA complexes in many aspects. Firstly, PfTFIIB-DNA complexes show decreased mobility in comparison to PfTBP-DNA complex. Secondly, the binding activity of PfTFIIB is lower than that of PfTBP and the PfTFIIB-DNA complexes are more sensitive to challenges with nonspecific competitor than PfTBP-DNA complexes. Thirdly, PfTFIIB binds *P. falciparum* promoter regions with higher affinity than the model Ad2ML promoter, whereas PfTBP binds both promoters with similar affinity. These differences in binding affinity, complex stability and mobility collectively suggest different DNA-binding characteristics of PfTBP and PfTFIIB.

The extensive PfTBP-independent binding activity of PfTFIIB reported in this study, has not been detected previously in EMSAs with other eukaryotic TFIIBs. Eukaryotic TFIIB binds TBP and stabilises the resulting TBP-TFIIB-DNA complex via interactions with BREs. Considering this known DNA-binding activity of eukaryotic TFIIB via BREs, it is interesting to speculate if the PfTFIIB interactions with DNA observed in this study are akin to the specific interactions with BREs observed in other eukaryotes. Eukaryotic TFIIB contacts the upstream IIB recognition sequence, BREu, within the DNA major groove, upstream of the TATA-box via helix-turn-helix (HTH) motif, a trihelical bundle in the second direct repeat of TFIIB core domain (Lagrange et al., 1998, Tsai and Sigler, 2000). TFIIB is known to interact with BREu in a TBP-independent manner. However, these interactions were determined by crystallographic and photo-crosslinking studies and have not been previously detected in EMSA conditions (see Chapter 1.2.4). The downstream IIB recognition sequence, BREd, is distinct from BREu and its TBP-dependent interaction with TFIIB is mediated by TFIIB recognition loop within the first direct repeat (Deng and Roberts, 2005). Sequence analysis and structural predictions of PfTFIIB indicate a notable conservation in the BREd recognition region of the first direct repeat (Steven Bing, MSc thesis 2015). However, similarly to plants and fungi (Lagrange et al., 1998), only a few basic residues of the HTH motif are conserved in PfTFIIB (Steven Bing, MSc thesis 2015). Crystallographic and mutational studies are necessary for any conclusive evidence regarding the conservation of function of these binding motifs in *P. falciparum* TFIIB. BREs are G/C-rich sequences (BREu: 5’-G/C-G/C-G/A-C-G-C-C-3’; BREd: 5’-G/A-T-T/G/A-T/G-G/ T-T/G-T/ G-3’) and show relatively low conservation within eukaryotes (Gershenzon and Ioshikhes, 2005). In archaea, despite the presence of the highly conserved HTH motif (Kosa et al., 1997),

the BREu sequence has low homology to eukaryotic BREs and is notably A/T-rich (Qureshi and Jackson, 1998).

The A/T-rich *P. falciparum* promoter regions used in this study do not contain the consensus BREd or BREu sequences, whereas the Ad2ML promoter contains both a BREu and a 6/7 match to the consensus BREd. However, our preliminary data indicate a higher binding affinity of PfTFIIB towards the *P. falciparum kahrp* promoter than Ad2ML. Given the low homology between human and *Plasmodium* TFIIB and the uniquely A/T-rich genome of *P. falciparum*, the extensive PfTBP-independent DNA-binding could be, at least in part, explained by significant alteration of the BRE consensus sequences in *P. falciparum* and alternative DNA-binding motifs of PfTFIIB.

In native polyacrylamide EMSAs, PfTFIIBΔ140 and PfTFIIBΔ37 formed distinct complexes with 60bp *kahrp* and *var* probes. Multiple complexes were seen with the *var* probe and could indicate the presence of multiple binding sites or a conformational change of DNA induced by PfTFIIB. To date, there is no evidence for TFIIB-induced DNA-bending in eukaryotic or archaeal systems in the absence of TBP. In Mg^{2+} agarose EMSAs, increasing PfTFIIB concentrations further decreased the PfTFIIB-DNA shift, indicating successive binding of multiple PfTFIIB molecules, that bind to lower affinity binding sites upon increasing PfTFIIB concentrations. Therefore, the multiple complexes seen in native PAGE EMSAs with *var* probe are likely due to multiple PfTFIIB molecules bound to the probe.

Similarly to PfTBP, PfTFIIB DNA-binding activity was increased at a lower temperature and resulting complexes were significantly more stable. The effect of incubation temperature on the activity of *P. falciparum* proteins was first noticed in this study and further assessments within our research group have thus far confirmed the preference for lower incubation temperatures also with other *P. falciparum* general transcription factors. The physiological environment of the parasite might be highly relevant in these sensitive assays and further studies are needed to investigate this possibility.

Eukaryotic TBP DNA-binding in then absence of TAFs or GTFs is unstable in EMSA conditions, however, our results suggest stable PfTBP-DNA complex formation at multiple locations. Considering this apparent promiscuity of PfTBP, it is possible, that PfTFIIB has an

extensive role in stabilisation of specific PfTBP-DNA complexes through promoter recognition via currently undescribed DNA-binding mechanisms and promoter elements. This is an interesting possibility that needs to be further investigated by PfTFIIB structural studies and binding site selection experiments.

4.7. Recruitment of PfTFIIB to promoter DNA by PfTBP

TFIIB plays a fundamental role in the stabilisation of the TBP-DNA complex in eukaryotes. Upon TBP binding to the TATA-motif, TFIIB binds beneath the TBP-TATA complex, where TFIIB first direct repeat of the core domain contacts TBP via its acidic carboy-terminal stirrup (Nikolov et al., 1995). As a result, TBP-TFIIB-promoter ternary complex is formed (Orphanides et al., 1996), which is stabilised by TFIIB sequence-specific contacts with BREs. TFIIB increases the affinity of TBP for the TATA box (Imbalzano, 1994), whereby the C-terminal TFIIB core domain is sufficient for the specific interactions with TBP and TBP-TFIIB-DNA complex formation (Nikolov et al., 1995). In this study, we report recruitment of PfTFIIB to the promoter by the presence of PfTBP under conditions, where independent PfTFIIB DNA-binding is suppressed. In agreement with the current eukaryotic model, PfTFIIB C-terminal core domain was found sufficient for the recruitment by PfTBP. We did not observe any significant stimulatory effect on PfTBP DNA-binding by the presence of PfTFIIB in immobilised template assays. Our findings provide first evidence for interactions between PfTBP and PfTFIIB and recruitment of PfTFIIB to promoter DNA by PfTBP.

4.8. PfTBP-PfTFIIB-promoter ternary complex formation

The stabilisation of promoter-bound TBP by TFIIB is thought to be caused by the combination of protein-protein interactions between TBP and TFIIB, and TFIIB interactions with BREs up- and downstream of the TATA-box. As a result, TBP-TFIIB-promoter ternary complex with an extended sequence specificity is formed (Imbalzano, 1994, Deng and Roberts, 2007). It has been shown that TBP-induced DNA-bending ensures the optimal spatial arrangement for TFIIB to access both BREs simultaneously (Lee and Hahn, 1995, Lagrange et al., 1998). Interestingly, our results indicate that PfTBP might not induce the degree of DNA-bending observed with

other eukaryotic TBPs due to the absence of a crucial Phe residue. The ability of TFIIA and TFIIB to recover mutated TBP with impaired DNA binding properties has been reported previously, where the addition of TFIIA or TFIIB results in TBP-TFIIA(TFIIB)-DNA complexes comparable to those observed with wild-type TBP (Lee et al., 1992, Kim et al., 1994).

The PfTFIIB and PfTBP orthologues utilised in this study possess strong DNA-binding activity in EMSA conditions, in contrast to their eukaryotic counterparts. In order to detect cooperative DNA-binding effect or stimulation of PfTBP DNA-binding in the presence of PfTFIIB, the independent DNA-binding of both protein factors was completely suppressed with poly(dG-dC). Incubation of PfTBP in the presence of PfTFIIB resulted in the formation of a distinct complex under these conditions. The mobility of the resulting complex was comparable to that of individual PfTBP- and PfTFIIB-DNA complexes. Thus, the observed complex is likely in an unbent conformation.

We tested for the presence of both PfTBP and PfTFIIB in the observed complex using custom-made polyclonal α-PfTBP and α-PfTFIIB antibodies. Inclusion of antibodies in the binding reaction often results in a further decrease of the ternary complex mobility, called a supershift. Alternatively, antibodies may inhibit protein-DNA complex formation and/or disrupt protein-DNA complexes. Indeed, inhibition of complex formation in the presence of specific antibody, rather than a supershift, has been observed previously with human TBP-TFIIB-promoter complex (Maldonado et al., 1990). Similarly, we observe complete inhibition of PfTBP-PfTFIIB-promoter complex in the presence of α-PfTBP or α-PfTFIIB antibodies, but not in the presence of equivalent amounts of α-PfTFIIA antibody. This inhibition is likely due to the antibodies obstructing PfTBP/PfTFIIB or PfTBP/DNA interaction surfaces. These results confirm the presence of both PfTBP and PfTFIIB in the complex. The formation of a novel complex with both protein factors present, under conditions where individual PfTBP and PfTFIIB DNA-binding is suppressed, provides strong evidence for the formation of a ternary PfTBP-PfTFIIB-promoter complex. Considering the novel DNA-binding characteristic of PfTFIIB observed in this study and the likely unbent conformation of the ternary complex, it is plausible that PfTFIIB stabilises the ternary complex by currently undescribed mechanisms. The DNA-binding experiments conducted in this study do not allow to infer the precise mechanism of the cooperativity between PfTBP and PfTFIIB. Crystallographic and structural studies are

necessary to further characterise PfTBP-PfTFIIB-promoter ternary complex.

Initial assessment of DNA sequence preference of the ternary complex was carried out with pre-assembled PfTBP-PfTFIIB-*kahrp* promoter complexes, that were challenged with unlabelled *kahrp, gbp-130* or Ad2ML promoter regions. Interestingly, no significant effect was seen with excess unlabelled Ad2ML promoter, whereas both *P. falciparum* putative promoter regions efficiently competed with the labelled *kahrp* promoter for the ternary complex binding. Previous experiments with isolated PfTBP indicated similar DNA-binding affinity on *P. falciparum* and Ad2ML promoters, whereas experiments in the presence of PfTFIIB results reveal a clearly higher preference of the PfTBP-PfTFIIB-promoter ternary complex for *P. falciparum* promoters. These results are in agreement with the model of extended sequence specificity of TBP-TFIIB-promoter ternary complex in comparison to TBP-promoter complex (Imbalzano, 1994). Therefore, it seems plausible, that PfTFIIB has developed mechanisms for extended sequence recognition in an unbent PfTBP-PfTFIIB-DNA complex. To our knowledge, this is the first demonstration of cooperative binding between *P. falciparum* general transcription factors and the formation of PfTBP-PfTFIIB-promoter ternary complex.

Conclusion

This MSc research reports the successful expression of PfTBPc and PfTFIIB. Expression vectors for wild-type and mutant PfTFIIB were constructed and standard purification scheme was established for all *P. falciparum* proteins used in this study. Initial characterisation of PfTBP and PfTFIIB revealed extensive DNA-binding activity of PfTBP and to date undescribed PfTBP-independent binding activity of PfTFIIB. Furthermore, to our knowledge, this is the first report of PfTBP and PfTFIIB interactions and ternary complex formation. Initial sequence preference for A/T-rich sequences was observed, with increased specificity of the ternary complex. This study provides an important starting point for the identification of the DNA elements directing transcription events in *P. falciparum* on a path towards understanding the complex gene expression regulation in this important parasite.

APPENDIX A

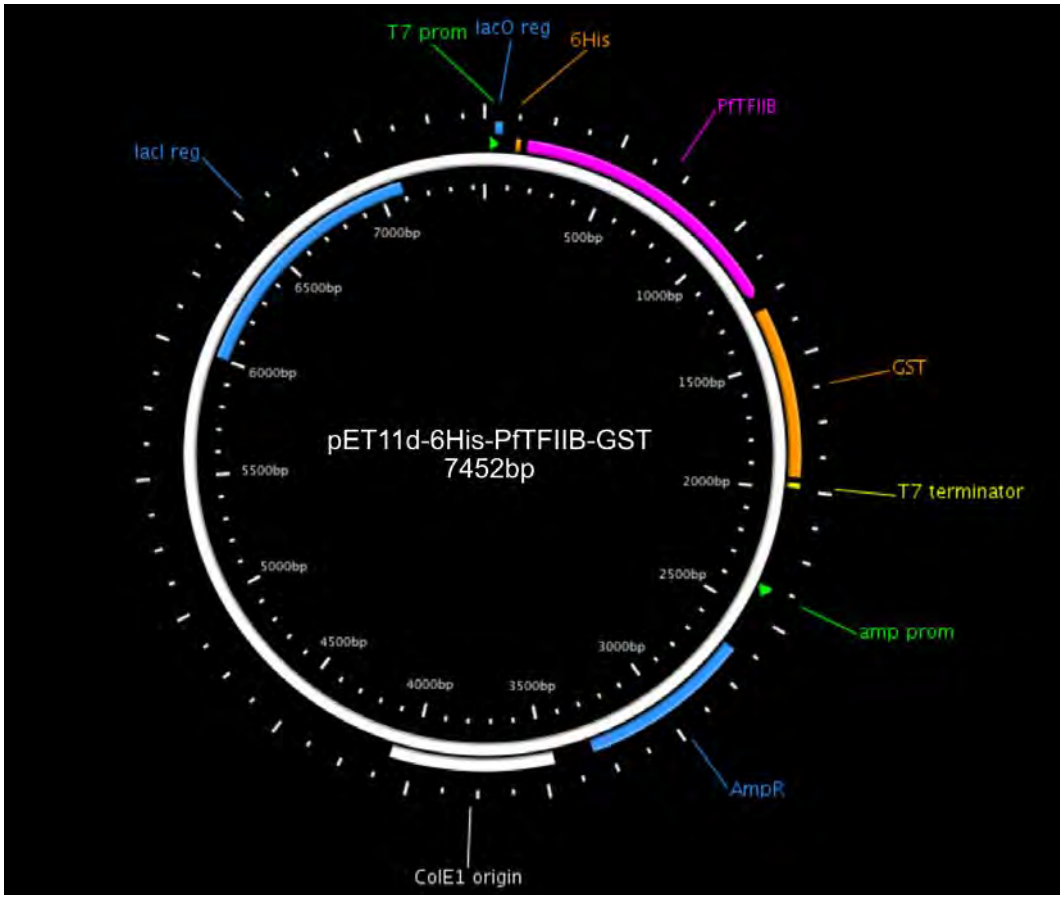

Figure 27. Plasmid map of pET11d-6His-PfTFIIB-GST

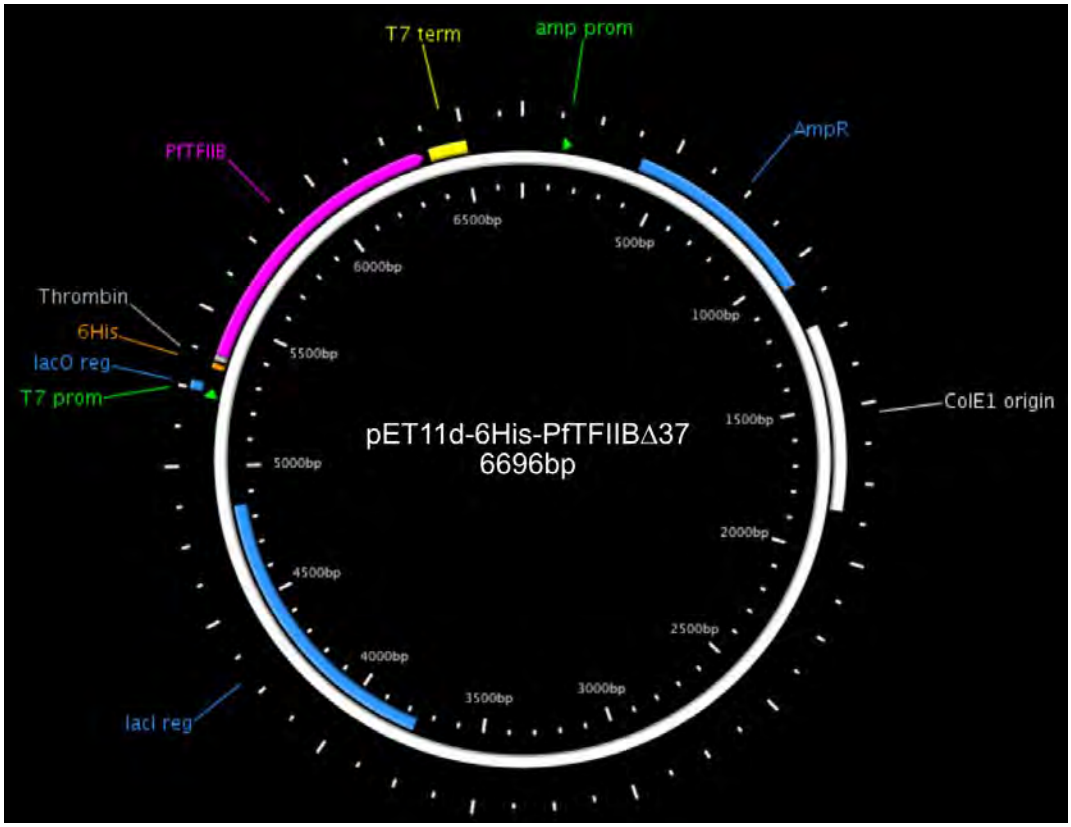

Figure 28. Plasmid map of pET11d-6His-PfTFIIBΔ37

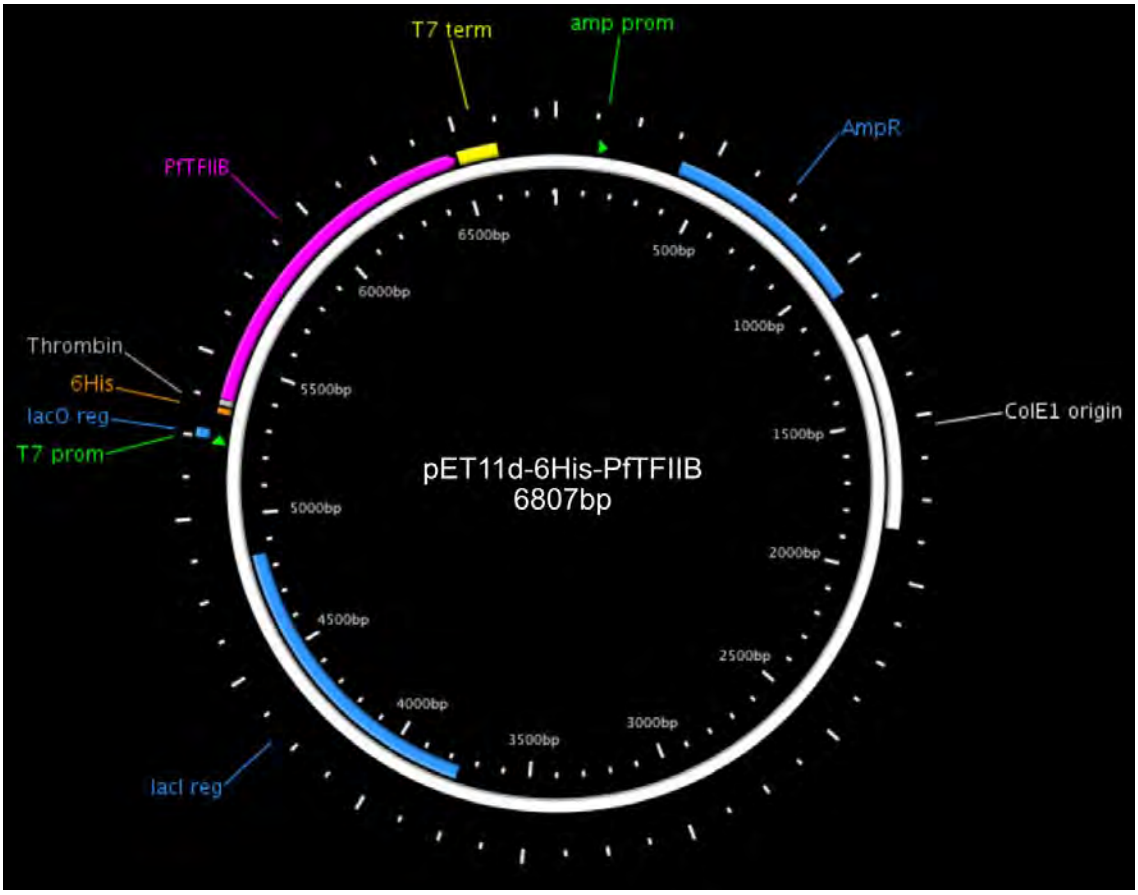

Figure 29. Plasmid map of pET11d-6His-PfTFIIB

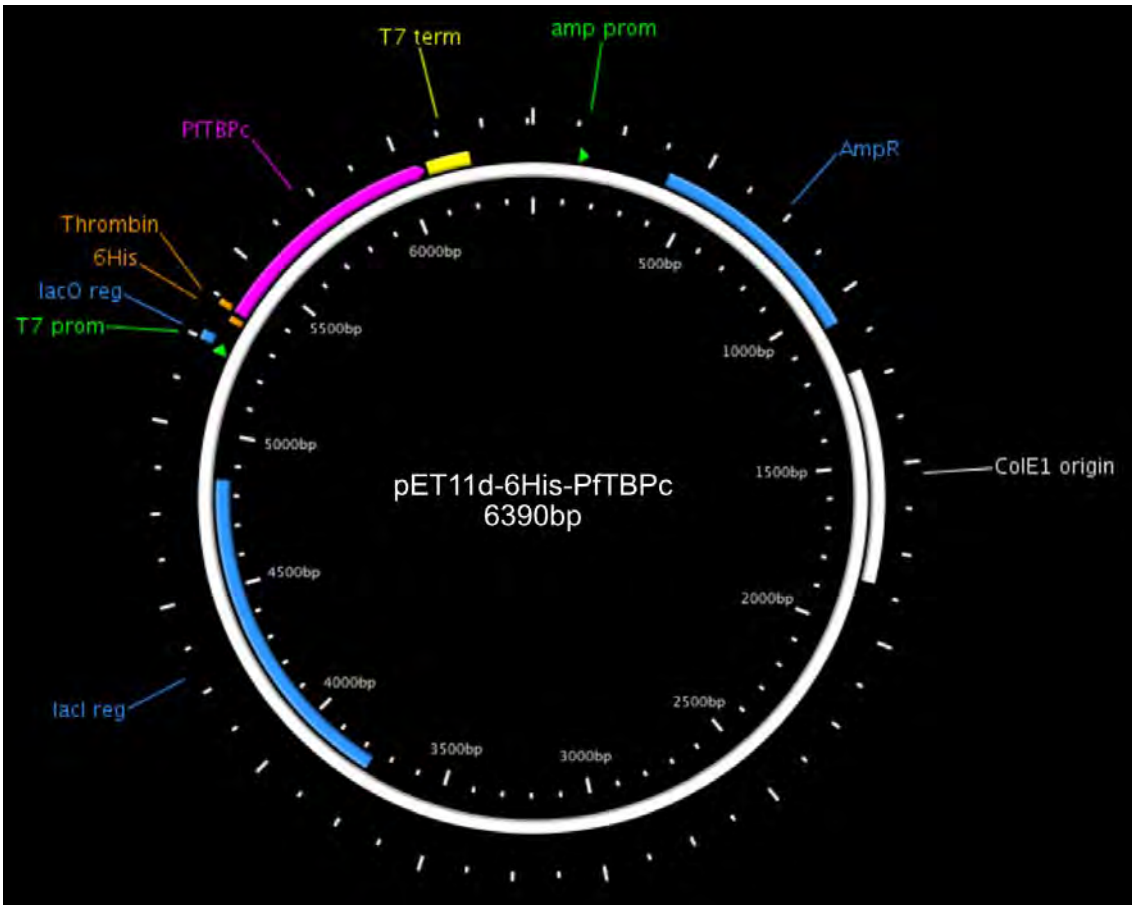

Figure 30. Plasmid map of pET11d-6His-PfTBPc

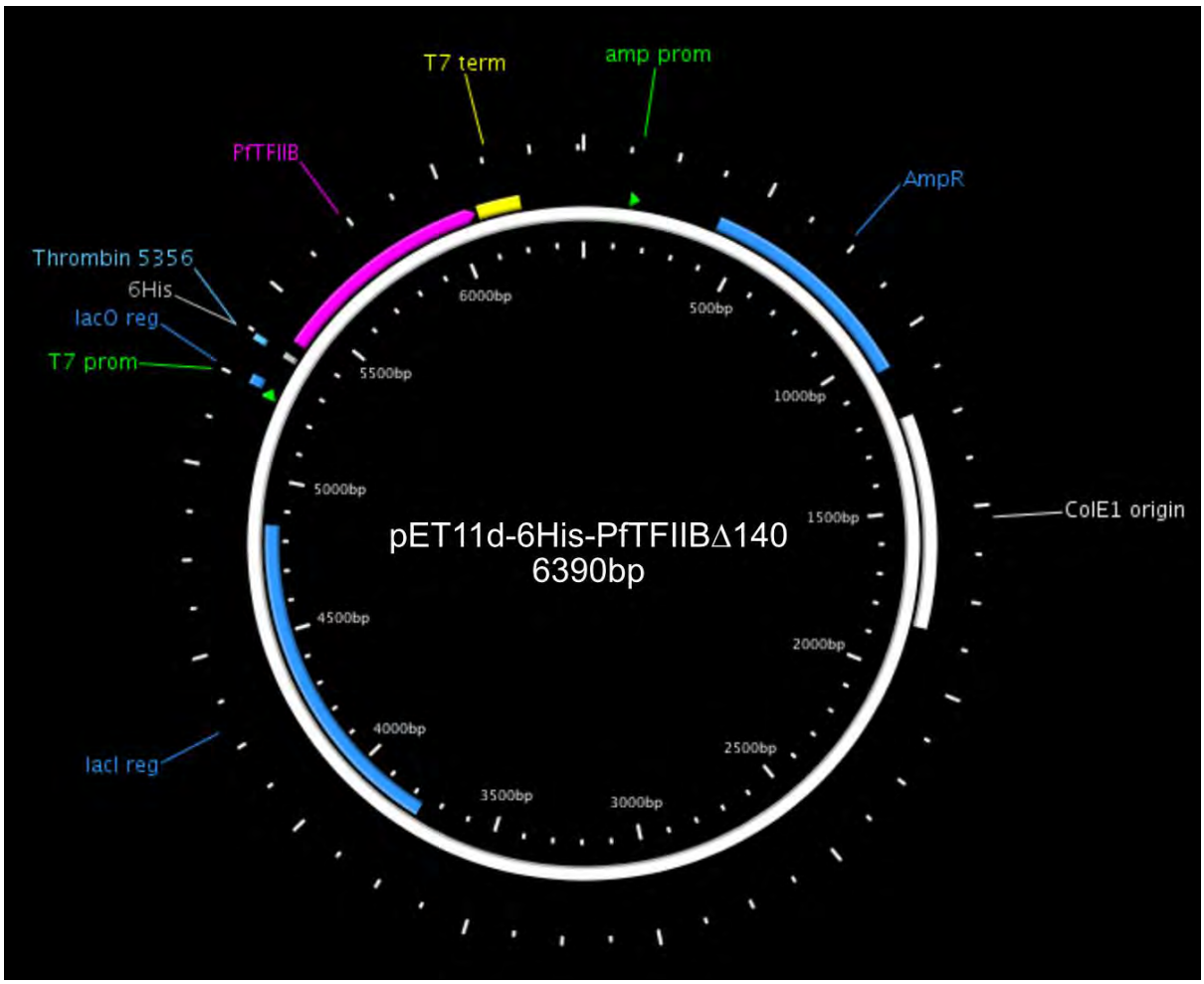

Figure 31. Plasmid map of pET11d-6His-PfTFIIBΔ140

APPENDIX B

Primer #	Primer	5'to 3' sequence	RE site
	Construction of pET11d-6His-PfTFIIB-GST		
1	T7T-FWD	5'-CACCGCTGAGCAAT AACTAGCATAA-3'	
2	*Eco*RI-REV	5'-CGTCTTCAAGAATTCTCATGTTTGACA-3'	*Eco*RI
3	BamHI-TEV-GST-FWD	5'-TTAAGATCTGGATCCGAAAACCTGTATTTTCAGGGCAT GTCCCCTATACTAGGTTATTGGAAA-3'	*Bam*HI
4	T7T-GST-REV	5'-TTATGCTAGTTATTGCTC AGCGGTGGCAGCAGCCAACCTATTATTTTGGAGGATGGT CGCCACCA-3'	
5	BamHI-TEV-FWD (used for overlap extension PCR)	5'-TTAAGATCTGGATCCGAAAACCTGTAT-3'	
6	*Eco*RI-REV (used for overlap extension PCR)	5'-CGTCTTCAAGAATTCTCATGTTTGACA-3'	
7	PfIIB-FWD (used for cDNA amplification)	5'-GGCAGCCATATGTCATCCATTTCTCCAAATAAAAAG-3'	*Nde*I
8	PfIIB-REV-2 (used for cDNA amplification)	5'-GCAGCCAGATCTGTTTTTTTTCCTGTCTTCACTAAG-3'	*Bgl*II
Construction of pET11d-6His-PfTFIIB; pET11d-6His-PfTFIIBΔ37; pET11d-6His-PfTFIIBΔ140			
9	PfIIB-FWD	5'-GGCAGCCATATGTCATCCATTTCTCCAAATAAAAAG-3'	*Nde*I
10	PfIIB-REV	5'-GCAGCCAGATCTCTATTAGTTTTTTTTC CTTTTGTCTTCACTAAG-3'	*Bgl*II
11	PfIIBΔ37-FWD	5'- GGCAGCCATATGAATGATCAAACATTA ATTTCCGCATTTAATATC-3'	*Nde*I
12	PfIIBΔ140-FWD	5'- GGCAGCCATATGTTAAGAAACTTTAGGAATAGTAGAAA-3'	*Nde*I

References

2015. *World Malaria Report* [Online]. World Health Organisation. Available: http://www.who.int/malaria/publications/world-malaria-report-2015/report/en/ [Accessed].

ALANO, P. 2007. Plasmodium falciparum gametocytes: still many secrets of a hidden life. *Mol Microbiol,* 66**,** 291-302.

BARUCH, D. I., PASLOSKE, B. L., SINGH, H. B., BI, X., MA, X. C., FELDMAN, M., TARASCHI, T. F. & HOWARD, R. J. 1995. Cloning the P. falciparum gene encoding PfEMP1, a malarial variant antigen and adherence receptor on the surface of parasitized human erythrocytes. *Cell,* 82**,** 77-87.

BIRKHOLTZ, L. M., BLATCH, G., COETZER, T. L., HOPPE, H. C., HUMAN, E., MORRIS, E. J., NGCETE, Z., OLDFIELD, L., ROTH, R., SHONHAI, A., STEPHENS, L. & LOUW, A. I. 2008. Heterologous expression of plasmodial proteins for structural studies and functional annotation. *Malar J,* 7**,** 197.

BISCHOFF, E. & VAQUERO, C. 2010. In silico and biological survey of transcription-associated proteins implicated in the transcriptional machinery during the erythrocytic development of Plasmodium falciparum. *BMC Genomics,* 11**,** 34.

BOLANOS-GARCIA, V. M. & DAVIES, O. R. 2006. Structural analysis and classification of native proteins from E. coli commonly co-purified by immobilised metal affinity chromatography. *Biochim Biophys Acta,* 1760**,** 1304-13.

BRANCUCCI, N. M., WITMER, K., SCHMID, C. D., FLUECK, C. & VOSS, T. S. 2012. Identification of a cis-acting DNA-protein interaction implicated in singular var gene choice in Plasmodium falciparum. *Cell Microbiol,* 14**,** 1836-48.

BUENDÍA-OROZCO, J. G., A.; PASTOR, N. 2005. Model of the TBP-TFIIB complex from Plasmodium falciparum: interface analysis and perspectives as a new target for antimalarial design. *Arch Med Res,* 36**,** 317-30.

BURATOWSKI, S., HAHN, S., GUARENTE, L. & SHARP, P. A. 1989. Five intermediate complexes in transcription initiation by RNA polymerase II. *Cell,* 56**,** 549-61.

BURLEY, S. K. & ROEDER, R. G. 1996. Biochemistry and structural biology of transcription factor IID (TFIID). *Annu Rev Biochem,* 65**,** 769-99.

BUSHNELL, D. A., WESTOVER, K. D., DAVIS, R. E. & KORNBERG, R. D. 2004. Structural basis of transcription: an RNA polymerase II-TFIIB cocrystal at 4.5 angstroms. *Science,* 303**,** 983-8.

CALLEBAUT, I., PRAT, K., MEURICE, E., MORNON, J. P. & TOMAVO, S. 2005. Prediction of the general transcription factors associated with RNA polymerase II in Plasmodium falciparum: conserved features and differences relative to other eukaryotes. *BMC Genomics,* 6**,** 100.

CAMPBELL, K. M. R., R.T.; STARGELL, L.A; LUMB, K.J; 2000. 2000 Reevaluation of transcriptional regulation by TATA-binding protein oligomerization: predominance of monomers. *Biochemistry,* 39**,** 2633-8.

CHAKRABARTI, D., REDDY, G. R., DAME, J. B., ALMIRA, E. C., LAIPIS, P. J., FERL, R. J., YANG, T. P., ROWE, T. C. & SCHUSTER, S. M. 1994. Analysis of expressed sequence tags from Plasmodium falciparum. *Mol Biochem Parasitol,* 66**,** 97-104.

CHALKLEY, G. E. & VERRIJZER, C. P. 1999. DNA binding site selection by RNA polymerase II TAFs: a TAF(II)250-TAF(II)150 complex recognizes the initiator. *EMBO J,* 18**,** 4835-45.

CHEN, H. T. & HAHN, S. 2004. Mapping the location of TFIIB within the RNA polymerase II transcription preinitiation complex: a model for the structure of the PIC. *Cell,* 119**,** 169-80.

CHEN, H. T., LEGAULT, P., GLUSHKA, J., OMICHINSKI, J. G. & SCOTT, R. A. 2000. Structure of a (Cys3His) zinc ribbon, a ubiquitous motif in archaeal and eucaryal transcription. *Protein Sci,* 9**,** 1743-52.

CHEN, Z. & MANLEY, J. L. 2003. Core promoter elements and TAFs contribute to the diversity of transcriptional activation in vertebrates. *Mol Cell Biol,* 23**,** 7350-62.

COLEMAN, R. A. & PUGH, B. F. 1995. Evidence for functional binding and stable sliding of the TATA binding protein on nonspecific DNA. *J Biol Chem,* 270**,** 13850-9.

COLEMAN, R. A., TAGGART, A. K., BURMA, S., CHICCA, J. J., 2ND & PUGH, B. F. 1999. TFIIA regulates TBP and TFIID dimers. *Mol Cell,* 4**,** 451-7.

COULSON, R. M., HALL, N. & OUZOUNIS, C. A. 2004. Comparative genomics of transcriptional control in the human malaria parasite Plasmodium falciparum. *Genome Res,* 14**,** 1548-54.

COWMAN, A. F., BERRY, D. & BAUM, J. 2012. The cellular and molecular basis for malaria parasite invasion of the human red blood cell. *J Cell Biol,* 198**,** 961-71.

COWMAN, A. F. & CRABB, B. S. 2006. Invasion of red blood cells by malaria parasites. *Cell,* 124**,** 755-66.

CRABB, B. S. & COWMAN, A. F. 1996. Characterization of promoters and stable transfection by homologous and nonhomologous recombination in Plasmodium falciparum. *Proc Natl Acad Sci U S A,* 93**,** 7289-94.

DE SILVA, E. K., GEHRKE, A. R., OLSZEWSKI, K., LEON, I., CHAHAL, J. S., BULYK, M. L. & LLINAS, M. 2008. Specific DNA-binding by apicomplexan AP2 transcription factors. *Proc Natl Acad Sci U S A,* 105**,** 8393-8.

DENG, W. & ROBERTS, S. G. 2005. A core promoter element downstream of the TATA box that is recognized by TFIIB. *Genes Dev,* 19**,** 2418-23.

DENG, W. & ROBERTS, S. G. 2007. TFIIB and the regulation of transcription by RNA polymerase II. *Chromosoma,* 116**,** 417-29.

FISHBURN, J. & HAHN, S. 2012. Architecture of the yeast RNA polymerase II open complex and regulation of activity by TFIIF. *Mol Cell Biol,* 32**,** 12-25.

FRIED, M. & DUFFY, P. E. 1996. Adherence of Plasmodium falciparum to chondroitin sulfate A in the human placenta. *Science,* 272**,** 1502-4.

GANGLOFF, Y. G., ROMIER, C., THUAULT, S., WERTEN, S. & DAVIDSON, I. 2001. The histone fold is a key structural motif of transcription factor TFIID. *Trends Biochem Sci,* 26**,** 250-7.

GARDNER, M. J., HALL, N., FUNG, E., WHITE, O., BERRIMAN, M., HYMAN, R. W., CARLTON, J. M., PAIN, A., NELSON, K. E., BOWMAN, S., PAULSEN, I. T., JAMES, K., EISEN, J. A., RUTHERFORD, K., SALZBERG, S. L., CRAIG, A., KYES, S., CHAN, M. S., NENE, V., SHALLOM, S. J., SUH, B., PETERSON, J., ANGIUOLI, S., PERTEA, M., ALLEN, J., SELENGUT, J., HAFT, D., MATHER, M. W., VAIDYA, A. B., MARTIN, D. M., FAIRLAMB, A. H., FRAUNHOLZ, M. J., ROOS, D. S., RALPH, S. A., MCFADDEN, G. I., CUMMINGS, L. M., SUBRAMANIAN, G. M., MUNGALL, C., VENTER, J. C., CARUCCI, D. J., HOFFMAN, S. L., NEWBOLD, C., DAVIS, R. W., FRASER, C. M. & BARRELL, B. 2002. Genome sequence of the human malaria parasite Plasmodium falciparum. *Nature,* 419**,** 498-511.

GARNER, M. M. & REVZIN, A. 1981. A gel electrophoresis method for quantifying the binding of proteins to specific DNA regions: application to components of the Escherichia coli lactose operon regulatory system. *Nucleic Acids Res,* 9**,** 3047-60.

GE, H., MARTINEZ, E., CHIANG, C. M. & ROEDER, R. G. 1996. Activator-dependent transcription by mammalian RNA polymerase II: in vitro reconstitution with general transcription factors and cofactors. *Methods Enzymol,* 274**,** 57-71.

GERSHENZON, N. I. & IOSHIKHES, I. P. 2005. Synergy of human Pol II core promoter elements revealed by statistical sequence analysis. *Bioinformatics,* 21**,** 1295-300.

GIETL, A., HOLZMEISTER, P., BLOMBACH, F., SCHULZ, S., VON VOITHENBERG, L. V., LAMB, D. C., WERNER, F., TINNEFELD, P. & GROHMANN, D. 2014. Eukaryotic and archaeal TBP and TFB/TF(II)B follow different promoter DNA bending pathways. *Nucleic Acids Res,* 42**,** 6219-31.

GISSOT, M. B., S.; REFOUR, P.; BOSCHET, C.; VAQUERO, C. 2005. PfMyb1, a Plasmodium falciparum Transcription Factor, is Required for Intra-erythrocytic Growth and Controls Key Genes for Cell Cycle Regulation. *Journal of Molecular Biology***,** 29-42.

GOPPELT, A. & MEISTERERNST, M. 1996. Characterization of the basal inhibitor of class II transcription NC2 from Saccharomyces cerevisiae. *Nucleic Acids Res,* 24**,** 4450-5.

GRUNBERG, S. & HAHN, S. 2013. Structural insights into transcription initiation by RNA polymerase II. *Trends Biochem Sci,* 38**,** 603-11.

GUNASEKERA, A. M. M., A.; MILITELLO, K.T.; SIMS, J. S.; DONG, C.K.; GIERAHN, T.; LE ROCH, K.; WINZELER, E.; WIRTH, D. F. 2007. Regulatory motifs uncovered among gene expression clusters in Plasmodium falciparum. *Mol Biochem Parasitol***,** 19-30.

HAHN, S. & YOUNG, E. T. 2011. Transcriptional regulation in Saccharomyces cerevisiae: transcription factor regulation and function, mechanisms of initiation, and roles of activators and coactivators. *Genetics,* 189**,** 705-36.

HARPER, S. & SPEICHER, D. W. 2011. Purification of proteins fused to glutathione S-transferase. *Methods Mol Biol,* 681**,** 259-80.

HECKMAN, K. L. & PEASE, L. R. 2007. Gene splicing and mutagenesis by PCR-driven overlap extension. *Nat Protoc,* 2**,** 924-32.

HELLMAN, L. M. & FRIED, M. G. 2007. Electrophoretic mobility shift assay (EMSA) for detecting protein-nucleic acid interactions. *Nat Protoc,* 2**,** 1849-61.

HOOPES, B. C., LEBLANC, J. F. & HAWLEY, D. K. 1992. Kinetic analysis of yeast TFIID-TATA box complex formation suggests a multi-step pathway. *J Biol Chem,* 267**,** 11539-47.

HORIKOSHI, M., YAMAMOTO, T., OHKUMA, Y., WEIL, P. A. & ROEDER, R. G. 1990. Analysis of structure-function relationships of yeast TATA box binding factor TFIID. *Cell,* 61**,** 1171-8.

HORROCKS, P., WONG, E., RUSSELL, K. & EMES, R. D. 2009. Control of gene expression in Plasmodium falciparum - ten years on. *Mol Biochem Parasitol,* 164**,** 9-25.

IMBALZANO, A. N. Z., K. S.; KINGSTON, R.E. 1994. Transcription factor (TF) IIB and TFIIA can independently increase the affinity of the TATA-binding protein for DNA. *J Biol Chem,* 269**,** 8280-6.

JOHNSON, K. M., WANG, J., SMALLWOOD, A. & CAREY, M. 2004. The immobilized template assay for measuring cooperativity in eukaryotic transcription complex assembly. *Methods Enzymol,* 380**,** 207-19.

JUVEN-GERSHON, T. & KADONAGA, J. T. 2010. Regulation of gene expression via the core promoter and the basal transcriptional machinery. *Dev Biol,* 339**,** 225-9.

KIM, J. L., NIKOLOV, D. B. & BURLEY, S. K. 1993. Co-crystal structure of TBP recognizing the minor groove of a TATA element. *Nature,* 365**,** 520-7.

KIM, T. H., BARRERA, L. O., ZHENG, M., QU, C., SINGER, M. A., RICHMOND, T. A., WU, Y., GREEN, R. D. & REN, B. 2005. A high-resolution map of active promoters in the human genome. *Nature,* 436**,** 876-80.

KIM, T. K., HASHIMOTO, S., KELLEHER, R. J., 3RD, FLANAGAN, P. M., KORNBERG, R. D., HORIKOSHI, M. & ROEDER, R. G. 1994. Effects of activation-defective TBP mutations on transcription initiation in yeast. *Nature,* 369**,** 252-5.

KIRKMAN, L. A. & DEITSCH, K. W. 2012. Antigenic variation and the generation of diversity in malaria parasites. *Curr Opin Microbiol,* 15**,** 456-62.

KLEJMAN, M. P., ZHAO, X., VAN SCHAIK, F. M., HERR, W. & TIMMERS, H. T. 2005. Mutational analysis of BTAF1–TBP interaction/ BTAF1 can rescue DNA-binding defective TBP mutants. *Nucleic Acids Res,* 33**,** 5426-36.

KOKUBO, T., SWANSON, M. J., NISHIKAWA, J. I., HINNEBUSCH, A. G. & NAKATANI, Y. 1998. The yeast TAF145 inhibitory domain and TFIIA competitively bind to TATA-binding protein. *Mol Cell Biol,* 18**,** 1003-12.

KOKUBO, T. G., D. W.; WOOTTON, J. C.; HORIKOSHI, M.; ROEDER, R. G.; NAKATANI, Y. 1994. Molecular cloning of Drosophila TFIID subunits. *Nature,* 367(6462)**,** 484-7.

KOMAKI-YASUDA, K., OKUWAKI, M., NAGATA, K., KAWAZU, S. & KANO, S. 2013. Identification of a novel and unique transcription factor in the intraerythrocytic stage of Plasmodium falciparum. *PLoS One,* 8**,** e74701.

KOSA, P. F., GHOSH, G., DEDECKER, B. S. & SIGLER, P. B. 1997. The 2.1-Å crystal structure of an archaeal preinitiation complex: TATA-box-binding protein/transcription factor (II)B core/ TATA-box. *Proc Natl Acad Sci,* 94**,** 6042-7.

KOSTREWA, D., ZELLER, M. E., ARMACHE, K. J., SEIZL, M., LEIKE, K., THOMM, M. & CRAMER, P. 2009. RNA polymerase II-TFIIB structure and mechanism of transcription initiation. *Nature,* 462**,** 323-30.

KUDDUS, R. & SCHMIDT, M. C. 1993. Effect of the non-conserved N-terminus on the DNA binding activity of the yeast TATA binding protein. *Nucleic Acids Res,* 21**,** 1789-96.

LAGRANGE, T., KAPANIDIS, A. N., TANG, H., REINBERG, D. & EBRIGHT, R. H. 1998. New core promoter element in RNA polymerase II-dependent transcription: sequence-specific DNA binding by transcription factor IIB. *Genes Dev,* 12**,** 34-44.

LANZER, M. D. B., D.; RAVETCH, J. V. 1992a. A sequence element associated with the Plasmodium falciparum kahrp gene is the site of developmentally regulated protein–DNA interactions. *Nucleic Acids Research,* 20**,** 3051-3056.

LANZER, M. D. B., D.; RAVETCH, J. V. 1992b. Transcription mapping of a 100 kb locus of Plasmodium falciparum identifies an intergenic region in which transcription terminates and reinitiates. *EMBO J,* 11**,** 1949-1955.

LEE, D. K., DEJONG, J., HASHIMOTO, S., HORIKOSHI, M. & ROEDER, R. G. 1992. TFIIA induces conformational changes in TFIID via interactions with the basic repeat. *Mol Cell Biol,* 12**,** 5189-96.

LEE, S. & HAHN, S. 1995. Model for binding of transcription factor TFIIB to the TBP-DNA complex. *Nature,* 376**,** 609-12.

LI, W. B., BZIK, D. J., GU, H. M., TANAKA, M., FOX, B. A. & INSELBURG, J. 1989. An enlarged largest subunit of Plasmodium falciparum RNA polymerase II defines conserved and variable RNA polymerase domains. *Nucleic Acids Res,* 17**,** 9621-36.

LIEBERMAN, P. M. & BERK, A. J. 1994. A mechanism for TAFs in transcriptional activation: activation domain enhancement of TFIID-TFIIA--promoter DNA complex formation. *Genes Dev,* 8**,** 995-1006.

LIEBERMAN, P. M., SCHMIDT, M. C., KAO, C. C. & BERK, A. J. 1991. Two distinct domains in the yeast transcription factor IID and evidence for a TATA box-induced conformational change. *Mol Cell Biol,* 11**,** 63-74.

LINDSTROM, M. S., DEISENROTH, C. & ZHANG, Y. 2014. Putting a Finger on Growth Surveillance: Insight into MDM2 Zinc Finger-Ribosomal Protein Interactions. *Cell Cycle,* 6**,** 434-437.

LIU, X., BUSHNELL, D. A., WANG, D., CALERO, G. & KORNBERG, R. D. 2010. Structure of an RNA polymerase II-TFIIB complex and the transcription initiation mechanism. *Science,* 327**,** 206-9.

MALDONADO, E., HA, I., CORTES, P., WEIS, L. & REINBERG, D. 1990. Factors involved in specific transcription by mammalian RNA polymerase II: role of transcription factors IIA, IID, and IIB during formation of a transcription-competent complex. *Mol Cell Biol,* 10**,** 6335-47.

MALECOVA, B., CAPUTO, V. S., LEE, D. F., HSIEH, J. J. & OELGESCHLAGER, T. 2015. Taspase1 processing alters TFIIA cofactor properties in the regulation of TFIID. *Transcription,* 6**,** 21-32.

MALECOVA, B., GROSS, P., BOYER-GUITTAUT, M., YAVUZ, S. & OELGESCHLAGER, T. 2007. The initiator core promoter element antagonizes repression of TATA-directed transcription by negative cofactor NC2. *J Biol Chem,* 282**,** 24767-76.

MCANDREW, M. B., READ, M., SIMS, P. F. & HYDE, J. E. 1993. Characterisation of the gene encoding an unusually divergent TATA-binding protein (TBP) from the extremely A+T-rich human malaria parasite Plasmodium falciparum. *Gene,* 124**,** 165-71.

MILLER, L. H., BARUCH, D. I., MARSH, K. & DOUMBO, O. K. 2002. The pathogenic basis of malaria. *Nature,* 415**,** 673-9.

MONTGOMERY, J., MPHANDE, F. A., BERRIMAN, M., PAIN, A., ROGERSON, S. J., TAYLOR, T. E., MOLYNEUX, M. E. & CRAIG, A. 2007. Differential var gene expression in the organs of patients dying of falciparum malaria. *Mol Microbiol,* 65**,** 959-67.

NIKOLOV, D. B. & BURLEY, S. K. 1994. 2.1 A resolution refined structure of a TATA box-binding protein (TBP). *Nat Struct Biol,* 1**,** 621-37.

NIKOLOV, D. B., CHEN, H., HALAY, E. D., HOFFMAN, A., ROEDER, R. G. & BURLEY, S. K. 1996. Crystal structure of a human TATA box-binding protein/TATA element complex. *Proc Natl Acad Sci U S A,* 93**,** 4862-7.

NIKOLOV, D. B., CHEN, H., HALAY, E. D., USHEVA, A. A., HISATAKE, K., LEE, D. K., ROEDER, R. G. & BURLEY, S. K. 1995. Crystal structure of a TFIIB-TBP-TATA-element ternary complex. *Nature,* 377**,** 119-28.

OELGESCHLAGER, T., TAO, Y., KANG, Y. K. & ROEDER, R. G. 1998. Transcription activation via enhanced preinitiation complex assembly in a human cell-free system lacking TAFIIs. *Mol Cell,* 1**,** 925-31.

ORPHANIDES, G., LAGRANGE, T. & REINBERG, D. 1996. The general transcription factors of RNA polymerase II. *Genes & Development,* 10**,** 2657-2683.

PAINTER, H. J., CAMPBELL, T. L. & LLINAS, M. 2011. The Apicomplexan AP2 family: integral factors regulating Plasmodium development. *Mol Biochem Parasitol,* 176**,** 1-7.

PATIKOGLOU, G. A. K., J.L.; SUN, L.; YANG, S.H.; KODADEK, T.; BURLEY, S. K. 1999. TATA element recognition by the TATA box-binding protein has been conserved throughout evolution. *Genes & Development,* 13**,** 3217-30.

PEI, X., AN, X., GUO, X., TARNAWSKI, M., COPPEL, R. & MOHANDAS, N. 2005. Structural and functional studies of interaction between Plasmodium falciparum knob-associated histidine-rich protein (KAHRP) and erythrocyte spectrin. *J Biol Chem,* 280**,** 31166-71.

PERKINS, M. E. 1984. Surface proteins of Plasmodium falciparum merozoites binding to the erythrocyte receptor, glycophorin. *J Exp Med,* 160**,** 788-98.

PETRI, V., HSIEH, M. & BRENOWITZ, M. 1995. Thermodynamic and kinetic characterization of the binding of the TATA binding protein to the adenovirus E4 promoter. *Biochemistry,* 34**,** 9977-84.

QING, G., MA, L. C., KHORCHID, A., SWAPNA, G. V., MAL, T. K., TAKAYAMA, M. M., XIA, B., PHADTARE, S., KE, H., ACTON, T., MONTELIONE, G. T., IKURA, M. & INOUYE, M. 2004. Cold-shock induced high-yield protein production in Escherichia coli. *Nat Biotechnol,* 22**,** 877-82.

QURESHI, S. A. & JACKSON, S. P. 1998. Sequence-specific DNA binding by the S. shibatae TFIIB homolog, TFB, and its effect on promoter strength. *Mol Cell,* 1**,** 389-400.

RODGERS, J. T., PATEL, P., HENNES, J. L., BOLOGNIA, S. L. & MASCOTTI, D. P. 2000. Use of biotin-labeled nucleic acids for protein purification and agarose-based chemiluminescent electromobility shift assays. *Anal Biochem,* 277**,** 254-9.

ROEDER, R. G. 1998. Role of general and gene-specific cofactors in the regulation of eukaryotic transcription. *Cold Spring Harb Symp Quant Biol,* 63**,** 201-18.

RUVALCABA-SALAZAR, O. K., DEL CARMEN RAMIREZ-ESTUDILLO, M., MONTIEL-CONDADO, D., RECILLAS-TARGA, F., VARGAS, M. & HERNANDEZ-RIVAS, R. 2005. Recombinant and native Plasmodium falciparum TATA-binding-protein binds to a specific TATA box element in promoter regions. *Mol Biochem Parasitol,* 140**,** 183-96.

SAMBROOK, J. & RUSSELL, D. W. 2001. Molecular Cloning: A Laboratory Manual. *Cold spring harbor laboratory press*.

SCHERF, A., LOPEZ-RUBIO, J. J. & RIVIERE, L. 2008. Antigenic variation in Plasmodium falciparum. *Annu Rev Microbiol,* 62**,** 445-70.

SINHA, A., HUGHES, K. R., MODRZYNSKA, K. K., OTTO, T. D., PFANDER, C., DICKENS, N. J., RELIGA, A. A., BUSHELL, E., GRAHAM, A. L., CAMERON, R., KAFSACK, B. F., WILLIAMS, A. E., LLINAS, M., BERRIMAN, M., BILLKER, O. & WATERS, A. P. 2014. A cascade of DNA-binding proteins for sexual commitment and development in Plasmodium. *Nature,* 507**,** 253-7.

STARR, D. B., HOOPES, B. C. & HAWLEY, D. K. 1995. DNA Bending is an Important Component of Site-specific Recognition by the TATA Binding Protein. *J Mol Biol,* 250**,** 434-46.

SU, X. Z., WU, Y., SIFRI, C. D. & WELLEMS, T. E. 1996. Reduced extension temperatures required for PCR amplification of extremely A+T-rich DNA. *Nucleic Acids Res,* 24**,** 1574-5.

THOMAS, M. C. & CHIANG, C. M. 2006. The general transcription machinery and general cofactors. *Crit Rev Biochem Mol Biol,* 41**,** 105-78.

TRAN, K. & GRALLA, J. D. 2008. Control of the timing of promoter escape and RNA catalysis by the transcription factor IIb fingertip. *J Biol Chem,* 283**,** 15665-71.

TSAI, F. T. & SIGLER, P. B. 2000. Structural basis of preinitiation complex assembly on human pol II promoters. *EMBO J,* 19**,** 25-36.

VEDADI, M., LEW, J., ARTZ, J., AMANI, M., ZHAO, Y., DONG, A., WASNEY, G. A., GAO, M., HILLS, T., BROKX, S., QIU, W., SHARMA, S., DIASSITI, A., ALAM, Z., MELONE, M., MULICHAK, A., WERNIMONT, A., BRAY, J., LOPPNAU, P., PLOTNIKOVA, O., NEWBERRY, K., SUNDARARAJAN, E., HOUSTON, S., WALKER, J., TEMPEL, W., BOCHKAREV, A., KOZIERADZKI, I., EDWARDS, A., ARROWSMITH, C., ROOS, D., KAIN, K. & HUI, R. 2007. Genome-scale protein expression and structural biology of Plasmodium falciparum and related Apicomplexan organisms. *Mol Biochem Parasitol,* 151**,** 100-10.

VOSS, T. S., THOMPSON, J. K., WATERKEYN, J., FELGER, I., WEISS, N., COWMAN, A. F. & BECK, H. P. 2000. Genomic distribution and functional characterisation of two distinct and conserved Plasmodium falciparum var gene 5' flanking sequences. *Mol Biochem Parasitol,* 107**,** 103-15.

WAKAGURI, H., SUZUKI, Y., SASAKI, M., SUGANO, S. & WATANABE, J. 2009. Inconsistencies of genome annotations in apicomplexan parasites revealed by 5'-end-one-pass and full-length sequences of oligo-capped cDNAs. *BMC Genomics,* 10**,** 312.

WEIDEMAN, C. A., NETTER, R. C., BENJAMIN, L. R., MCALLISTER, J. J., SCHMIEDEKAMP, L. A., COLEMAN, R. A. & PUGH, B. F. 1997. Dynamic interplay of TFIIA, TBP and TATA DNA. *J Mol Biol,* 271**,** 61-75.

WENG, H., GUO, X., PAPOIN, J., WANG, J., COPPEL, R., MOHANDAS, N. & AN, X. 2014. Interaction of Plasmodium falciparum knob-associated histidine-rich protein (KAHRP) with erythrocyte ankyrin R is required for its attachment to the erythrocyte membrane. *Biochim Biophys Acta,* 1838**,** 185-92.

WONG, J. M. & BATEMAN, E. 1994. TBP-DNA interactions in the minor groove discriminate between A:T and T:A base pairs. *Nucleic Acids Res,* 22**,** 1890-6.

YOUNG, J. A., JOHNSON, J. R., BENNER, C., YAN, S. F., CHEN, K., LE ROCH, K. G., ZHOU, Y. & WINZELER, E. A. 2008. In silico discovery of transcription regulatory elements in Plasmodium falciparum. *BMC Genomics,* 9**,** 70.

YUDA, M., IWANAGA, S., SHIGENOBU, S., KATO, T. & KANEKO, I. 2010. Transcription factor AP2-Sp and its target genes in malarial sporozoites. *Mol Microbiol,* 75**,** 854-63.

ZERBY, D. & LIEBERMAN, P. M. 1997. Functional analysis of TFIID-activator interaction by magnesium-agarose gel electrophoresis. *Methods,* 12**,** 217-23.

ZHANG, D. Y., CARSON, D. J. & MA, J. 2002. The role of TFIIB-RNA polymerase II interaction in start site selection in yeast cells. *Nucleic Acids Res,* 30**,** 3078-85.

ZHANG, Z. & DIETRICH, F. S. 2005. Mapping of transcription start sites in Saccharomyces cerevisiae using 5' SAGE. *Nucleic Acids Res,* 33**,** 2838-51.

ZHAO, X. & HERR, W. 2002. A regulated two-step mechanism of TBP binding to DNA: a solvent-exposed surface of TBP inhibits TATA box recognition. *Cell,* 108**,** 615-27.